

**Entwicklung und Validierung eines Vignettentests
zur Erfassung des fachdidaktischen Wissens im
Leseunterricht bei angehenden Lehrkräften**

Von der Pädagogischen Hochschule Heidelberg
zur Erlangung des Grades einer Doktorin der Philosophie (Dr. phil.)

genehmigte Dissertation von

Juliane Rutsch, geb. Kiermeier

aus

Sinsheim, Baden-Württemberg

2016

Erstgutachter:	Prof. Dr. Tobias Dörfler
Zweitgutachterin:	Prof. Dr. Cornelia Glaser
Fach:	Psychologie
Tag der Mündlichen Prüfung:	06.12.2016

Abstract

In der vorliegenden Studie wurde ein standardisiertes Testverfahren zur Erfassung des fachdidaktischen Wissens im Leseunterricht bei angehenden Lehrkräften entwickelt und validiert. Dazu wurden basierend auf der lesedidaktischen und kognitionspsychologischen Forschung Testaufgaben in Form von Unterrichtssituationen (sog. Vignetten) erstellt, die von angehenden Lehrkräften hinsichtlich einer fachdidaktisch relevanten Fragestellung in einem geschlossenen Antwortformat bewertet werden sollten.

Die inhaltliche Validität der Testaufgaben wurde durch eine mehrstufige Expertenbefragung mit qualitativen ($N = 7$) und quantitativen ($N = 87$) Befragungen sichergestellt.

Zur Überprüfung der psychometrischen Eigenschaften wurden Deutschstudierende aller Semester und Referendare mit dem Fach Deutsch ($N = 581$) sowie Kontrollgruppen (fachfremdes Lehramt, Germanistik, Psychologie; $N = 147$) untersucht.

Die Analysen zur Datenstruktur zeigen den guten Modellfit eines eindimensionalen Modells des lesedidaktischen Wissens (χ^2 ($N = 578$, $df = 54$) = 73.67, $p = 0.04$, $\chi^2/df = 1.36$, RMSEA = 0.03, CFI = 0.93). Der Vignettentest wurde anschließend mit dem Partial Credit Model skaliert. Die Infit- und Outfitindices bescheinigen eine sehr gute Passung der Vignetten zum Modell (Werte zwischen 0.98 und 1.00, alle $ps \leq 0.05$). Die Reliabilität des Testverfahrens ist als akzeptabel zu bezeichnen (EAP-Reliabilität = 0.59). Weiterführende Analysen zur Validität des Testverfahrens haben gezeigt, dass sowohl die Lehrerfahrung als auch die Nutzung spezifischer Lerngelegenheiten die Testleistung der Studierenden signifikant vorhersagen (alle $ps \leq 0.05$). Zudem weisen Deutschstudierende einen signifikant höheren Testscore als die Psychologiestudierenden und die Germanistikstudierenden ($p \leq 0.05$) und einen deskriptiv höheren Testscore als die Studierenden fachfremder Lehrämter auf.

Diese Ergebnisse sollen durch anschließende Studien mit teilweise längsschnittlichen Fragestellungen und Designs ergänzt und präzisiert werden.

Danksagung

Ich möchte mich bei folgenden Personen für Ihre Unterstützung, Begleitung und Beratung bedanken:

Herrn Prof. Dr. Tobias Dörfler danke ich für die Bereitstellung des äußerst interessanten Promotionsthemas, das die Entwicklung und Validierung eines standardisierten vignettenbasierten Testverfahrens zur Erfassung des fachdidaktischen Wissens im Leseunterricht bei angehenden Lehrkräften zum Ziel hatte. Ich verdanke ihm jede erdenkliche Unterstützung und viele hilfreiche inhaltliche sowie methodische Diskussionen. Die einzelnen Arbeitsschritte dieses Promotionsvorhabens wurden von ihm äußerst professionell, kompetent und intensiv begleitet. Vor allem bedanke ich mich für das mir entgegengebrachte Vertrauen und die gestalterische Freiheit bei der inhaltlichen Ausrichtung des Vignettentests. Insbesondere seine kritischen Anmerkungen und weiterführenden Gedanken zur ersten Version dieser Arbeit haben substantiell zu ihrer Qualität beigetragen.

Frau Prof. Dr. Cornelia Glaser danke ich für die Bereitschaft, meine Dissertation als Zweitgutachterin zu beurteilen. Auch sie hat mich im Laufe der letzten 2.5 Jahre im Rahmen verschiedener Diskussionen durch konstruktive Anregungen und Verbesserungsvorschläge äußerst kompetent unterstützt.

Herrn Dr. Markus Schmitt danke ich für die fachkundige und hilfreiche Unterstützung im Rahmen der Erstellung der Testaufgaben.

Frau Eva Keller danke ich für den kollegialen Austausch und die enge Zusammenarbeit als „Schwesterteilprojekte“. Durch ihre Fach- und Methodenkompetenz sowie ihre Zuverlässigkeit und ihr organisatorisches Geschick (insbesondere bei der Durchführung der Haupterhebung) wurde meine Arbeit maßgeblich bereichert.

Frau Julia Ollesch danke ich – als Bürokollegin – für die vielfältige Unterstützung und Diskussionen während der unterschiedlichen (Arbeits-)Phasen des Promotionskollegs.

Allen Kollegiaten sowie Betreuern des Forschungs- und Nachwuchskollegs „Effektive Kompetenzdiagnose in der Lehrerbildung“ (EKoL) danke ich für die kollegiale Zusammenarbeit und die vielen anregenden Diskussionen vor allem im Rahmen der verschiedenen Scientific Retreats und Arbeitstreffen.

Meinem Mann Christoph Rutsch danke ich für die zahlreichen Gespräche und Diskussionen bezüglich meines Promotionsthemas ebenso wie für die emotionale Unterstützung und sein unermüdliches Ermutigen, vor allem während schwieriger Arbeitsphasen. Außerdem danke ich ihm für die erste Rückmeldung zu dieser Dissertation, da er diese als Erster lesen und korrigieren „durfte“.

Meiner besten Freundin Alexandra Edinger danke ich ebenfalls für ihr aufmerksames und dezidiertes Korrekturlesen und die vielen inhaltlichen Anregungen, die meine Arbeit verbessert haben.

Ann-Kathrin Metzger danke ich für das gründliche und aufmerksame Korrekturlesen der fertiggestellten Arbeit.

Meiner Schwester Katrin Kiermeier danke ich für das kurzfristige, äußerst genaue und hilfreiche Korrekturlesen des Diskussionsteils sowie für die vielen hilfreichen und kritischen Formulierungsvorschläge und -anregungen.

Inhaltsübersicht

ABBILDUNGSVERZEICHNIS	10
TABELLENVERZEICHNIS	12
ABKÜRZUNGSVERZEICHNIS	13
1 EINLEITUNG UND ZIELSTELLUNG	14
2 PARADIGMEN DER PROFESSIONSFORSCHUNG ZUM LEHRERBERUF	17
3 PROFESSIONELLE KOMPETENZ VON LEHRKRÄFTEN	20
4 MODELL DER PROFESSIONELLEN KOMPETENZ VON LEHRKRÄFTEN	23
5 PROFESSIONELLES WISSEN VON LEHRKRÄFTEN	30
6 LESEKOMPETENZ UND LESEPROZESS	53
7 PROFESSIONELLES WISSEN VON (ANGEHENDEN) DEUTSCHLEHRKRÄFTEN	62
8 ABLEITUNG DER FORSCHUNGSANLIEGEN UND HYPOTHESEN	74
9 TESTENTWICKLUNG	84
10 METHODE	106
11 ERGEBNISSE	119
12 DISKUSSION	143
13 LIMITATIONEN DER VORLIEGENDEN STUDIE	164
14 IMPLIKATIONEN FÜR DIE PRAXIS	168
15 AUSBLICK ÜBER ANSCHLIEßENDE FORSCHUNGSFRAGEN	175
16 FAZIT	179
17 LITERATURVERZEICHNIS	FEHLER! TEXTMARKE NICHT DEFINIERT.
ANHANG: VIGNETTEN DES FINALEN TESTINSTRUMENTS	213

Inhaltsverzeichnis

ABBILDUNGSVERZEICHNIS	10
TABELLENVERZEICHNIS	12
ABKÜRZUNGSVERZEICHNIS	13
1 EINLEITUNG UND ZIELSTELLUNG	14
2 PARADIGMEN DER PROFESSIONSFORSCHUNG ZUM LEHRERBERUF	17
3 PROFESSIONELLE KOMPETENZ VON LEHRKRÄFTEN	20
4 MODELL DER PROFESSIONELLEN KOMPETENZ VON LEHRKRÄFTEN	23
4.1 Überzeugungen/Werthaltungen/Ziele.....	24
4.2 Motivationale Orientierung.....	25
4.3 Selbstregulation	28
5 PROFESSIONELLES WISSEN VON LEHRKRÄFTEN	30
5.1 Inhaltliche Kategorisierung des professionellen Wissens von Lehrkräften.....	30
5.2 Formale Kategorisierung des professionellen Wissens von Lehrkräften	33
5.3 Inhaltlich-formale Kategorisierung des professionellen Wissens von Lehrkräften.....	34
5.4 Zusammenhang zwischen professionellem Wissen und professionellem Handeln	36
5.5 Entwicklung des professionellen Wissensinhalts PCK bei (angehenden) Lehrkräften.....	40
5.5.1 Allgemeine Entwicklung von PCK.....	40
5.5.2 Einflussfaktoren auf die Entwicklung von PCK.....	43
5.6 Struktur des professionellen Wissensinhalts PCK	44
5.7 Testinstrumente zur Erfassung des professionellen Wissensinhalts PCK	48
6 LESEKOMPETENZ UND LESEPROZESS	53
6.1 Lesekompetenz: Begriff und Relevanz.....	53
6.2 Der Leseprozess	54
6.3 Empirische Evidenzen zur Lesekompetenz deutscher Schüler.....	57
6.4 Determinanten von Lesekompetenz	58
6.4.1 Basale Lesefertigkeiten	59
6.4.2 Lesemotivation.....	60
6.4.3 Lesestrategien	61

7	PROFESSIONELLES WISSEN VON (ANGEHENDEN) DEUTSCHLEHRKRÄFTEN	62
7.1	Fachdidaktisches Wissen im Leseunterricht	63
7.1.1	Fachdidaktische Ansätze zur Steigerung basaler Lesefertigkeiten	64
7.1.2	Fachdidaktische Ansätze zur Steigerung der Lesemotivation.....	66
7.1.3	Fachdidaktische Ansätze zu Förderung des Einsatzes von Lesestrategien	66
7.1.4	Diagnose von Lese- und Verstehensprozessen.....	68
7.1.5	Fazit zu fachdidaktischen Ansatzpunkten zur Förderung von Lesekompetenz	70
7.2	Empirische Evidenzen zum fachdidaktischen Wissen (angehender) Lehrkräfte im Leseunterricht	71
8	ABLEITUNG DER FORSCHUNGSANLIEGEN UND HYPOTHESEN	74
9	TESTENTWICKLUNG	84
9.1	Ableitung eines theoretischen Modells des lese-didaktischen Wissens für die Sekundarstufe I	84
9.2	Entwicklung der Testaufgaben	87
9.3	Inhaltliche Validierung der Vignetten und Erstellung des Lösungsmusters	90
9.3.1	Qualitative Expertenbefragung	91
9.3.2	Quantitative Expertenbefragung	92
9.3.3	Pilotierung	96
9.3.4	Generierung der Auswertungsnorm	100
10	METHODE	106
10.1	Stichprobe	106
10.2	Verlauf der Hauptstudie	107
10.3	Verwendete Instrumente	108
10.3.1	Soziodemographische Angaben	109
10.3.2	Vignettentest zum schreibdidaktischen Wissen	109
10.3.3	Abiturdurchschnitt und Sprachliches Denken.....	110
10.3.4	PK	111
10.3.5	Lesedidaktische Überzeugungen.....	111
10.3.6	Berufsbezogenes Selbstkonzept.....	112
10.3.7	Berufswahlmotive	112
10.3.8	BIG-5	113
10.4	Analysen	113
10.4.1	Umgang mit fehlenden Werten	115
10.4.2	Überprüfung der Voraussetzungen der Analysen	116
10.4.3	Hierarchische Datenstruktur	117
11	ERGEBNISSE	119
11.1	Deskriptive Analysen	119

11.2 Forschungsanliegen 1	122
11.2.1 Überprüfung der Datenstruktur	123
11.2.2 Skalierung	125
11.3 Forschungsanliegen 2	130
11.3.1 Hypothese 1: Kontrollgruppenvergleich	130
11.3.2 Hypothese 2: Anstieg des lesedidaktischen Wissens im Querschnitt	131
11.3.3 Hypothese 3: Einfluss der Lehrererfahrung auf das lesedidaktische Wissen	134
11.3.4 Hypothese 4: Einfluss lesedidaktischer Seminare auf das lesedidaktische Wissen	136
11.3.5 Hypothese 5: Praktische Erfahrung und Studiendauer als inkrementelle Prädiktoren des lesedidaktischen Wissens	137
11.4 Forschungsanliegen 3	138
11.4.1 Hypothese 6: Zusammenhang zu PK	140
11.4.2 Hypothese 7: Zusammenhang zu lesedidaktischen Überzeugungen	141
11.4.3 Hypothese 8: Zusammenhang zum berufsbezogenen Selbstkonzept	141
11.4.4 Hypothese 9: Zusammenhang zur Berufswahlmotivation	141
11.4.5 Hypothese 10: Zusammenhang zu den BIG-5	141
12 DISKUSSION	143
12.1 Inhaltsvalidität und Auswertungsnorm	143
12.2 Datenstruktur	144
12.3 Skalierung	147
12.4 Validität	150
12.4.1 Kontrollgruppenvergleich	150
12.4.2 Anstieg des lesedidaktischen Wissens über das Studium (im Querschnitt)	153
12.4.3 Bedeutung der praktischen Erfahrungen für die Entwicklung des lesedidaktischen Wissens	155
12.5 Zusammenhänge des lesedidaktischen Wissens zu verwandten Konstrukten	157
13 LIMITATIONEN DER VORLIEGENDEN STUDIE	164
14 IMPLIKATIONEN FÜR DIE PRAXIS	168
15 AUSBLICK ÜBER ANSCHLIEßENDE FORSCHUNGSFRAGEN	175
16 FAZIT	179
17 LITERATURVERZEICHNIS	FEHLER! TEXTMARKE NICHT DEFINIERT.
ANHANG: VIGNETTEN DES FINALEN TESTINSTRUMENTS	213

Abbildungsverzeichnis

Abbildung 1. Das Modell der professionellen Kompetenz aus dem Forschungsprogramm COACTIV. Eigene Darstellung nach Kunter et al. (2011, S. 32).	23
Abbildung 2. Konzepte des „Lehrerwissens“. Eigene und erweiterte Darstellung nach Neuweg (2014, S. 585).	38
Abbildung 3. Modell „Five Stages of Skill Acquisition“. Eigene und erweiterte Darstellung nach Dreyfus (2004, S. 181).	41
Abbildung 4. Vereinfachtes Modell des Textverstehens in Anlehnung an Kintsch (1992) und Artelt und Dörfler (2010).	56
Abbildung 5. Theoretisches Strukturmodell des lesedidaktischen Wissens.	85
Abbildung 6. Beispielvignette aus dem Testinstrument zur Facette „Diagnose/ ...	88
Abbildung 7. Darstellung des Prozesses der inhaltlichen Validierung der Testaufgaben.	91
Abbildung 8. Übersicht über die Institutionen, an denen die Experten tätig sind. Mehrfachnennungen sind möglich.	92
Abbildung 9. Darstellung der quadrierten Mahalanobis-Distanzen.	117
Abbildung 10. Beispiel einer bearbeiteten Vignette. In der Abbildung ist beispielhaft das Kreuzmuster eines Probanden dargestellt. Der Proband in diesem Beispiel hat Item 1 als die (laut Expertenmeinung) am schlechtesten geeignete Handlungsalternative erkannt.	120
Abbildung 11. Zweidimensionales Modell des lesedidaktischen Wissens mit den Facetten Förderwissen/Intervention und Diagnose/Inhaltswissen. ...	123
Abbildung 12. Eindimensionales Modell mit dem Faktor „Lesedidaktisches Wissen“.	124
Abbildung 13. Graphische Darstellung der Q3-Statistiken zwischen den Itemrelationen.	126
Abbildung 14. Graphische Darstellung der Q3-Statistiken zwischen den Vignetten.	126
Abbildung 15. Person-Item Map zur Darstellung der Verteilungen der Itemkategorieschwellen δ_k und der Personenfähigkeiten.	128
Abbildung 16. Wright Map zur Darstellung der Thurstonian Thresholds.	129
Abbildung 17. Einfaktorielle ANCOVA zum Vergleich zwischen den Kontrollgruppen	131
Abbildung 18. Regression des latenten Faktor Lesedidaktisches Wissen auf die Studiendauer (Semester 1 bis Referendariat).	132
Abbildung 19. t -Tests für die verschiedenen Substichproben mit den Gruppen „Beginn des Studiums“ (Semester 1-3) und „Ende des Studiums“ (Semester 4-8), * $p \leq 0.05$.	133
Abbildung 20. Einfaktorielle ANOVA (n. s.) zum Anstieg des lesedidaktischen Wissens über die Semester für die PH-Sekundarstufendeutschlehramtsstudierenden mit einfachen geplanten Kontrasten zum ersten Semester, * $p \leq 0.05$.	134

Abbildung 21. Regression des latenten Faktor Lesedidaktisches Wissen auf die Lehrerfahrung im Allgemeinen (ja/nein).	135
Abbildung 22. Regression des latenten Faktor Lesedidaktisches Wissen auf die Lehrerfahrung im Leseunterricht (ja/nein).....	135
Abbildung 23. Regression des latenten Faktor Lesedidaktisches Wissen auf „Besuch lesedidaktisches Seminar“ (ja/nein).	136
Abbildung 24. Latentes Regressionsmodell für die Regression des lesedidaktischen Wissens auf die Semesteranzahl und den latenten Faktor „Praktische Erfahrungen“	137
Abbildung 25. Wissensverlust bei Personalabgängen (aus: Sutter, 2016, S. 22).	171

Tabellenverzeichnis

Tabelle 1. Konzeptualisierung von PCK (eigene Darstellung nach Depaepe et al., 2013).	46
Tabelle 2. Zuordnung der Vignetten zu Modellfacette und lesedidaktischem Forschungsbereich	89
Tabelle 3. Einschätzung der Vignetten in der quantitativen Expertenbefragung...	95
Tabelle 4. Beurteilung der Vignetten im Studierendenpilot.....	97
Tabelle 5. Darstellung der Vignetten des finalen Testinstruments.....	100
Tabelle 6. Expertenratings der Vignette „Überprüfen des Textverständnisses“ .	101
Tabelle 7. Z-standardisierten Expertenratings der Vignette „Überprüfen des Textverständnisses“ und Median der Items	102
Tabelle 8. Z-standardisierte Ratings zweier Probanden als Beispiel	103
Tabelle 9. Vergleich der Experten-Rangfolge der Items mit der Rangfolge der Probanden	103
Tabelle 10. Punktvergabe für die Studierenden anhand des Vergleichs der Antwortrelationen der Experten und der Probanden	104
Tabelle 11. Beschreibung der verschiedenen Substichproben	107
Tabelle 12. Deskriptive Statistiken für die einzelnen Vignetten.....	122
Tabelle 13. Bivariate Korrelationen zwischen dem Testscore (lesedidaktisches Wissen) und verwandten Konstrukten	139
Tabelle 14. Partielle Korrelation des Testscores mit weiteren Konstrukten unter Kontrolle von Abiturdurchschnittsnote und Sprachlichem Denken .	140

Abkürzungsverzeichnis

CK	Content Knowledge (Fachwissen)
IRT	Item Response Theorie
KMK	Kultusministerkonferenz
MZP	Messzeitpunkt
PCK	Pedagogical Content Knowledge (Fachdidaktisches Wissen)
P(P)K	Pedagogical Knowledge (Pädagogisch psychologisches Wissen)

1 Einleitung und Zielstellung

Nationale und internationale Schulleistungsuntersuchungen haben gezeigt, dass vergleichsweise viele Kinder und Jugendliche in Deutschland eine eingeschränkte Lesekompetenz aufweisen (Bos, Tarelli, Bremerich-Vos & Schwippert, 2012; Hohn, Schiepe-Tiska, Sälzer & Artelt, 2013). Daher drängt sich von bildungs- und gesellschaftspolitischer Seite folgende Frage auf: Welche Ansatzpunkte und Maßnahmen können dazu beitragen, die Lesekompetenz von Schülern¹ nachhaltig und effektiv zu fördern?

Neben der Entwicklung und Evaluation von Fördermaßnahmen der Lesekompetenz auf Schülerebene (Rosebrock & Nix, 2014) steigt das Interesse daran zu untersuchen, welche Kompetenzen eine Lehrkraft eigentlich benötigt, um den Leseunterricht erfolgreich und effizient gestalten zu können. Nicht zuletzt deshalb, weil mittlerweile eine ausreichende Anzahl an empirischen Studien im bildungswissenschaftlichen Kontext gezeigt hat, dass Lehrkräfte die wichtigsten pädagogischen Akteure im Wirkungsbereich von Schule und ihrer Organisation sind (vgl. Fend, 2008; Schleicher, 2012).

Initiale Forschungsarbeiten im Fachbereich Mathematik liefern gesicherte empirische Evidenzen zum Professionalisierungsprozess angehender Lehrkräfte im Studium und Referendariat (Kunter, Baumert et al., 2011): Es wird im Allgemeinen davon ausgegangen, dass ein Bündel, bestehend aus verschiedenen kognitiven und affektiven Teilkompetenzen, die Lehrkräfte dazu befähigt, erfolgreich zu unterrichten. Professionelle Wissensinhalte von Lehrkräften wurden als bedeutsame Teilkompetenz intensiv beforscht, da insbesondere der professionelle Wissensinhalt *fachdidaktisches Wissen* einen maßgeblichen Anteil an der Unterrichtsqualität zu haben scheint (Voss, Kunina-Habenicht, Hoehne & Kunter, 2015).

Um gesicherte Aussagen über den Professionalisierungsprozess von (angehenden) Lehrkräften treffen zu können, müssen die Teilkompetenzen, die zum erfolgreichen Handeln im Unterricht notwendig sind, empirisch erfassbar gemacht werden.

¹ In der vorliegenden Arbeit wird im Sinne einer besseren Lesbarkeit das inkludierende Maskulinum verwendet. Dieses spricht Männer und Frauen selbstverständlich gleichermaßen an.

So wurden in jüngster Zeit verschiedene Erhebungsinstrumente u. a. zu professionellen Wissensinhalten (Tepner et al., 2012), Überzeugungen (Chan, 2011), Motivation (Pohlmann & Möller, 2010) und selbstregulativen Fähigkeiten (Schaarschmidt & Fischer, 1997) entwickelt und eingesetzt.

Nur wenn professionelle Kompetenzen und Wissensinhalte von Lehrkräften empirisch verlässlich zu erfassen sind, können evidenzbasiert Implikationen und Handlungsmöglichkeiten abgeleitet werden, die die Qualität in der Lehramtsausbildung sichern bzw. steigern können.

Wie aber läuft der Professionalisierungsprozess bei angehenden Deutschlehrkräften ab? Welche Lerngelegenheiten sind besonders bedeutsam, welche Möglichkeiten gibt es hier zur Optimierung? Hierzu liegt bisher eine eingeschränkte empirische Befundlage vor (Dehrmann & Standke, 2012; Herzmann & König, 2016). Es ist allerdings davon auszugehen, dass eine Kompetenzsteigerung auf Seiten der Lehrkräfte sich auf die Schülerebene niederschlägt (Harnischfeger & Wiley, 1977). Daher stellt die Untersuchung von professionellen Wissensinhalten von (angehenden) Lehrkräften bezogen auf den Leseunterricht ein bedeutsames Forschungsdesiderat dar.

Das Ziel der vorliegenden Studie ist daher, ein standardisiertes Testinstrument zu entwickeln und zu validieren, welches das fachdidaktische Wissen im Leseunterricht bei angehenden Lehrkräften zuverlässig erfassen kann. Es wird dazu ein Verfahren zur Erfassung professioneller Wissensinhalte von Lehrkräften vorgestellt und eingesetzt, der sog. Vignettentest. Ein Vignettentest ist ein Testformat, das aus spezifischen Testaufgaben – Vignetten – besteht. Vignetten sind kompakte Beschreibungen von authentischen Unterrichtssituationen, die von Probanden hinsichtlich fachdidaktisch relevanter Fragestellungen eingeschätzt bzw. bearbeitet werden sollen.

Das Herzstück dieser Dissertationsschrift bildet die Ableitung eines theoretischen Strukturmodells des fachdidaktischen Wissens im Leseunterricht aus der lesedidaktischen und kognitionspsychologischen Literatur sowie die darauf aufbauende Entwicklung von entsprechenden Unterrichtsvignetten. Nach der inhaltlichen Validierung der Vignetten wurde das Testinstrument in einer großflächigen Studierendenbefragung eingesetzt. Im Ergebnisteil werden sowohl Befunde zu den

psychometrischen Eigenschaften des Testinstruments als auch erste Erkenntnisse zu Struktur und (querschnittlicher) Entwicklung des fachdidaktischen Wissens im Leseunterricht über das Lehramtsstudium hinweg berichtet.

In der Diskussion sollen diese Ergebnisse dann unter Rückbezug auf die theoretischen Grundlagen der Studie ausführlich reflektiert werden. Einen weiteren zentralen Punkt dieser Dissertationsschrift bildet die Ableitung von Implikationen für die Praxis, die aus den erzielten Forschungsergebnissen resultieren.

Die Arbeit wird schließlich mit einem Ausblick über anschließende Forschungsfragen und -möglichkeiten sowie einem schlussfolgernden Fazit enden.

2 Paradigmen der Professionsforschung zum Lehrerberuf

Seit Mitte der 1990er Jahre werden in Deutschland große nationale und internationale Schulleistungsuntersuchungen durchgeführt (Bos, Tarelli et al., 2012; Bos, Wendt, Köller & Selter, 2012; Klieme et al., 2010). Die Ergebnisse dieser Studien haben eine breite öffentliche Diskussion zu Fragen des Bildungssystems in Deutschland angeregt, da den deutschen Schülern in verschiedenen Bereichen der schulischen Bildung im OECD-Vergleich mittelmäßige bis unterdurchschnittliche Leistungen bescheinigt wurden. In Folge dessen wurde damit begonnen, sich mit möglichen Ursachen für das enttäuschende Abschneiden sowie mit Ansatzpunkten zur Verbesserung der Leistungsentwicklung deutscher Schüler zu beschäftigen.

Verschiedene prominente empirische Forschungsarbeiten der letzten Jahre haben gezeigt, dass die Lehrtätigkeit einschließlich der Klassenführung einen bedeutsamen Einfluss auf die Schülerleistungsentwicklung hat (Hanushek & Rivkin, 2010; Hattie, 2009; Lipowsky, 2006). Die Lehrkraft nimmt im Unterrichtsgeschehen eine zentrale Rolle ein: Neben der Planung des Unterrichts ist sie für die Gestaltung einer lernförderlichen Lernatmosphäre verantwortlich (Lindmeier, 2012). Ein Ansatzpunkt zur Verbesserung der Schülerleistungen in allen Bereichen der schulischen Bildung stellt daher die intensive Beschäftigung mit professionellen Merkmalen von Lehrkräften dar. Dies führt zu einer zunehmend systematischen Untersuchung der Eigenschaften von Lehrkräften in Deutschland und einer Diskussion über die Qualität der Lehramtsausbildung. Daneben begann nach 1999 die Umsetzung der Bologna-Reform, die sich unter anderem ebenfalls zum Ziel setzte, die Ausbildung von Lehrkräften zu optimieren.

Die Frage, die sich allerdings stellt, ist: Was kann denn eigentlich unter einem „guten Lehrer“ verstanden werden? Was zeichnet ihn aus? Sicherlich können die meisten Personen, basierend auf eigenen Erfahrungen aus der Schulzeit, Lehrkräfte benennen, die sie als „gut“ empfunden haben. Doch welche Eigenschaften sind das, die den guten Lehrer auszeichnen? Ist er besonders nett und freundlich, streng, strukturiert oder motivierend? Es fällt schwer, hier eine konkrete, allge-

mein gültige Aussage zu treffen. So unterscheiden sich auch Schüler unterschiedlicher Schulformen in ihren Vorstellungen über „den guten Lehrer“: Gymnasias-ten empfinden vor allem pädagogisches Engagement als wichtige Eigenschaft von Lehrkräften, Nicht-Gymnasiasten betonen eher, dass das Durchsetzungsvermögen von Lehrkräften bedeutsam ist (König, 2007).

Die Frage nach dem guten Lehrer beschäftigt auch im wissenschaftlichen Kontext:

„Seit mehr als einem Jahrhundert beschäftigt sich die pädagogisch-psychologische Forschung mit der Persönlichkeit von Lehrerinnen und Lehrern, ihrem pädagogischen Handeln, mit der Bedeutung didaktischer Expertise und den Wirkungen des Unterrichts auf das Erleben, Verhalten und Lernen der Schülerinnen und Schüler. Was ist der gegenwärtige Erkenntnisstand dieser vielfältigen empirischen Forschungsbemühungen? Gibt es den ‚guten Lehrer‘, ‚die gute Lehrerin‘ überhaupt, und, wenn ja, wodurch lassen sie sich charakterisieren? Sind bestimmte Persönlichkeitsmerkmale entscheidend, spielen wirksame Lehrtechniken die dominierende Rolle oder geht es bevorzugt um die professionalisierte Unterrichtsperspektive?“ (Weinert, 1996, S. 141).

In diesem Zitat gibt Weinert einen prägnanten Überblick über die verschiedenen Forschungstraditionen und -ansätze, die sich zur Aufgabe gemacht haben die Kernessenz, die den „guten Lehrer“ ausmacht, herauszukristallisieren:

Zwischen 1900 und 1960 war das *Persönlichkeitsparadigma* in der Forschung zum Lehrerberuf vorherrschend. Dieser Forschungsansatz fokussiert vorrangig Persönlichkeitsmerkmale von Lehrkräften und versucht aufbauend darauf Schlussfolgerungen über die Unterrichtsqualität und die Schülerleistungen zu erhalten (Überblick: Pause, 1970). Im aktuellen Diskurs versteht man unter der Persönlichkeit von Lehrkräften ein „Ensemble relativ stabiler Dispositionen, die für das Handeln, den Erfolg und das Befinden im Lehrerberuf bedeutsam sind“ (Mayr & Neuweg, 2006, S. 183). Persönlichkeitseigenschaften von (angehenden) Lehrkräften stellen eine der (Lern-)Voraussetzungen zum Erwerb professioneller Kompetenzen dar, weshalb Persönlichkeitseigenschaften von Lehrkräften nach wie vor im Interesse der Forschung zur Professionalisierung im Lehrerberuf stehen (Bau-

mert & Kunter, 2006; Kunter, Kleickmann, Klusmann & Richter, 2011; Mayr, 2014).

Um 1960 rückte unter dem Einfluss des Behaviorismus zunehmend das *Prozess-Produkt-Paradigma* in den Vordergrund. In diesem Forschungsansatz wurden vorrangig die im Unterrichtsgeschehen ablaufenden *Prozesse* fokussiert. Man beschäftigte sich insbesondere mit dem Einfluss verschiedener Lehreraktivitäten auf Schüleroutcomevariablen wie Leistungsentwicklung oder Motivation. Da der Ansatz des Prozess-Produkt-Paradigmas sich auf den Wirkzusammenhang Lehrerverhalten-Schüleroutcome fokussierte, stand hier weniger die Person des Lehrers im Forschungszentrum, sondern viel mehr das Verhalten des Lehrers in bestimmten Unterrichtssituationen (Überblick: Brophy & Good, 1986). Im Rahmen der empirischen Forschung des Prozess-Produkt-Paradigmas wurden verschiedene lern- und leistungsrelevante Aspekte des unterrichtlichen Handelns von Lehrkräften identifiziert, beispielsweise die Prävention von Unterrichtsstörungen oder ein angemessenes Unterrichtstempo (vgl. Besser & Krauss, 2009).

Ab den 1980er Jahren wendete sich die empirische Bildungsforschung wieder zunehmend der Person des Lehrers zu und orientierte sich an Evidenzen aus der kognitionspsychologischen Forschung zum Expertenwissen. Bis heute ist das sogenannte *Experten-Paradigma* in der Forschung zum Lehrerberuf das zentrale. Hier stehen allerdings nicht, wie im Persönlichkeitsparadigma, verschiedene Persönlichkeitseigenschaften der Lehrkräfte im Vordergrund, sondern vielmehr die Wissensinhalte und Kompetenzen, die den Lehrer dazu befähigen, effizient unterrichten zu können. Dadurch, dass Kompetenzen und Wissensinhalte im Gegensatz zu bestimmten Persönlichkeitseigenschaften grundsätzlich als erlernbar gelten, wird durch die Hinwendung zum Experten-Paradigma im Lehrerberuf zunehmend über die Möglichkeiten der Entwicklung und Bedingungen des Erwerbs professioneller Kompetenzen nachgedacht. Darüber hinaus werden die Bedeutsamkeit von persönlichen Erfahrungen und Überzeugungssystemen, die sich auf Lehr-Lernprozesse beziehen, zur Entwicklung von professioneller Kompetenz betont.

Tatsächlich ist bei den Übergängen zwischen den unterschiedlichen Paradigmen in der Forschung zum Lehrerberuf nicht von echten Brüchen zu sprechen; es handelt sich vielmehr um eine Verschiebung von Akzenten (vgl. Baumert, Blum & Neubrand, 2004; Krauss & Bruckmeier, 2014).

3 Professionelle Kompetenz von Lehrkräften

In Zusammenhang mit den Überlegungen zum „guten Lehrer“ und der Entwicklung des Expertenparadigmas wurde und wird über die sog. „professionelle Kompetenz“ von Lehrkräften diskutiert. Der Terminus „professionelle Kompetenz“ setzt sich aus zwei Begrifflichkeiten zusammen: Der Begriff „Profession“ meint so viel wie Arbeit, Beruf oder Beschäftigung². Handelt eine Person professionell, also im Sinne ihrer Profession, hat sie Kenntnisse über allgemeingültige Regeln, die der Person helfen, einen gegebenen Sachverhalt zu überblicken, Handlungsalternativen zu generieren und sich nach diesen zu verhalten (Reinisch, 2009).

Der Begriff Kompetenz meint nach Weinert „die bei Individuen verfügbaren oder durch sie erlernbaren kognitiven Fähigkeiten und Fertigkeiten, um bestimmte Probleme zu lösen, sowie die damit verbundenen motivationalen, volitionalen und sozialen Bereitschaften und Fähigkeiten, um die Problemlösungen in variablen Situationen erfolgreich und verantwortungsvoll nutzen zu können“ (Weinert, 2001, S. 27). Der Kompetenzbegriff umfasst daher einen funktionalen (die Lehrperson *kann* eine professionelle Handlung ausführen) sowie einen volitionalen (die Lehrperson *will* eine professionelle Handlung ausführen) Aspekt (Vogelsang & Reinhold, 2013). Verfügt eine Person über berufsbezogene, professionelle Kompetenzen, dann verfügt sie gleichsam über das Handwerkszeug ihren Beruf erfolgreich ausüben zu können. Der Terminus „professionelle Kompetenz“ subsummiert folgerichtig die Voraussetzungen, über die Lehrkräfte verfügen sollten, um Unterricht erfolgreich gestalten zu können (Kunter & Trautwein, 2013).

Da (professionelle) Kompetenzen laut Weinert erlernbar sind, stellt sich die Frage, wie sich professionelle Kompetenzen von Lehrpersonen optimal entwickeln. In der Literatur werden verschiedene Annahmen zum Professionalisierungsprozess von Lehrpersonen diskutiert: Der strukturtheoretische Ansatz (Helsper, 1995,

² Die ursprüngliche Unterscheidung zwischen Professionen (klassische Professionen sind Ärzte, Anwälte, Kleriker) und gewöhnlichen Berufen spielt in der aktuellen Diskussion zur Professionalisierungsforschung eine äußerst geringe Rolle (Bonnet & Herniks, 2014) und ist daher zu vernachlässigen.

2014), der berufsbiographische Ansatz (Herzog, 2014; Terhart, 2007) und der kompetenztheoretische Ansatz (Baumert & Kunter, 2006).

Der strukturtheoretische Ansatz charakterisiert die berufsbezogenen Anforderungen und Aufgaben an Lehrpersonen als grundsätzlich widersprüchlich. Als Beispiel wäre hier das Spannungsfeld „Einheitlichkeit“ versus „Differenz“ zu nennen (formale Gleichstellung aller Schüler, aber sehr individuelle Entwicklungsverläufe, die u.U. individuelles Handeln erfordern). Professionalität im Lehrerberuf zeigt sich laut den Vertretern des strukturtheoretischen Ansatzes darin, dass Lehrpersonen mit Spannungen, die durch berufsbedingte Widersprüchlichkeiten erzeugt werden, adäquat umgehen können.

Der berufsbiographische Ansatz betont allem voran die berufsbiographische Entwicklung von Lehrpersonen. Hier wird eine dynamische und individuelle Perspektive auf die Entwicklung professioneller Kompetenzen im Lehrberuf eingenommen, die durch persönliche Lebensereignisse, Weiterbildungserfahrungen, Karrieremuster sowie Belastungserfahrungen und deren Bewältigung geprägt sind. In diesem Ansatz wird daher hervorgehoben, dass es sich bei der Lehrerausbildung nicht um einen auf das Studium und Referendariat begrenzten Zeitraum handelt, sondern dass Lehrer(aus)bildung das gesamte Berufsleben umfasst.

Der kompetenztheoretische Ansatz legt eine genaue Analyse der für den Lehrerberuf notwendigen Aufgaben und Anforderungen zugrunde (z.B. „Schülerleistungen korrekt diagnostizieren können“). Daraus werden bestimmte Kompetenzen und professionelle Wissensdomänen abgeleitet, die für das erfolgreiche Bewältigen dieser Aufgaben und Anforderungen notwendig sind. Zur Festlegung dieser Kompetenzen werden sowohl theoretische Überlegungen als auch empirische Evidenzen mit einbezogen. Dieser Ansatz ist innerhalb der pädagogischen Psychologie dem unter 2 dargestellten Experten-Paradigma zuzuordnen, das annimmt, dass für die Steuerung von Lehr- und Lernprozessen professionellen Wissensinhalten von Lehrkräften eine bedeutsame Rolle zukommt (Krauss & Bruckmeier, 2014).

Der kompetenztheoretische Ansatz erhebt außerdem – im Gegensatz zum strukturtheoretischen und berufsbiographischen Ansatz – den Anspruch, empirisch überprüfbar zu sein. Dieser Ansatz verfolgt das Ziel, Modelle zu Struktur und

Entwicklung von Handlungskompetenzen und Wissensinhalten, die für den Lehrerberuf bedeutsam sind, abzuleiten und empirisch zu prüfen.

Die vorliegende Arbeit legt den kompetenztheoretischen Ansatz der Entwicklung professioneller Kompetenzen von Lehrpersonen zugrunde. Im nachfolgenden Kapitel wird ein prominentes Modell des kompetenztheoretischen Ansatzes vorgestellt und diskutiert.

4 Modell der professionellen Kompetenz von Lehrkräften

Die Autoren des Forschungsprogramms COACTIV haben basierend auf dem kompetenztheoretischen Ansatz ein viel beachtetes strukturelles Modell der professionellen Kompetenz von Lehrkräften vorgelegt (Baumert & Kunter, 2011a). Das besondere Anliegen war hier „einen theoretischen und empirischen Beitrag zur Klärung zentraler Konzepte und zur Diskussion über die Professionalisierung von Lehrkräften [zu] leisten“ (Kunter, Baumert et al., 2011, S. 29). Die empirische Überprüfung des Modells erfolgte für Lehrkräfte und Lehramtsstudierende des Fachs Mathematik.

Das Kompetenzmodell von COACTIV besteht aus vier übergeordneten kognitiven und affektiven Aspekten professioneller Kompetenz, die für das erfolgreiche Handeln der Lehrkraft im Unterricht als notwendig erachtet werden (siehe Abbildung 1).

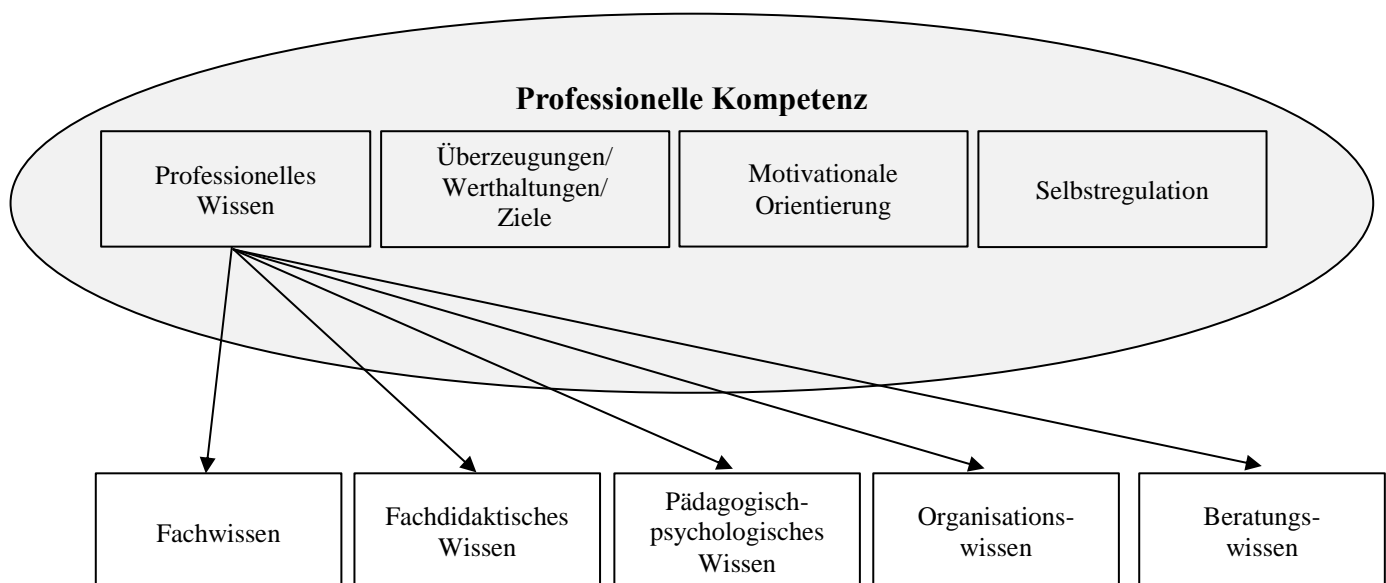


Abbildung 1. Das Modell der professionellen Kompetenz aus dem Forschungsprogramm COACTIV. Eigene Darstellung nach Kunter et al. (2011, S. 32).

Im Folgenden werden die vier übergeordneten Kompetenzaspekte des COACTIV-Modells erläutert sowie einschlägige empirische Forschungsbefunde dargestellt. Zunächst werden die Kompetenzaspekte *Überzeugungen/Werthaltungen/Ziele*, *Motivationale Orientierung* und *Selbstregulation* vorgestellt. Zuletzt wird ausführlich auf den Kompetenzbereich *Professionelles Wissen* eingegangen, da das Konzept des professionellen Wissens von Lehrkräften zentral für die vorliegende Dissertationsstudie ist.

4.1 Überzeugungen/Werthaltungen/Ziele

In dem Modell zur professionellen Kompetenz von Lehrkräften wird die Teilkompetenz „Überzeugungen/Werthaltungen/Ziele“ genannt. In der COACTIV-Studie werden Überzeugungen von Lehrkräften „definiert als überdauernde existentielle Annahmen über Phänomene oder Objekte der Welt, die subjektiv für wahr gehalten werden, sowohl implizite als auch explizite Anteile besitzen und die Art der Begegnung mit der Welt beeinflussen“ (Voss, Kleickmann, Kunter & Hachfeld, 2011, S. 235). Überzeugungen von Lehrkräften können verschiedene Bereiche des Lehrerberufs wie beispielsweise Lehr-Lernprozesse im Unterrichtsgeschehen oder auch das Bildungssystem als Ganzes betreffen. Es wird angenommen, dass Überzeugungen von Lehrkräften die Interaktionen mit den Schülern beeinflussen und so das Unterrichtsgeschehen vorstrukturieren können (Köller, Baumert & Neubrand, 2000). Des Weiteren können Überzeugungen von Lehrkräften einen Einfluss auf die Gestaltung des Unterrichts (operationalisiert über die Dimensionen *Effizienz der Klassenführung*, *Potenzial zur kognitiven Aktivierung* und *konstruktive Unterstützung*) und in Folge dessen auch auf Schülerleistungen nehmen (Voss et al., 2011).

In der Literatur werden u. a. in Bezug auf Lehr-Lernprozesse eine *konstruktivistische* und eine *transmissive* Überzeugung von Lehrkräften unterschieden (Behrmann & Souvignier, 2013; Chan, 2011; Dubberke, Kunter, McElvany, Brunner & Baumert, 2008). Unter einer konstruktivistischen Überzeugung wird eine Haltung von Lehrpersonen verstanden, die die Schüler als aktive Lernende versteht und weniger als passive Rezipienten von Lehrinhalten. Letzteres wird als eine transmissive Überzeugung verstanden. Es liegen empirische Evidenzen vor, dass eine konstruktivistische Überzeugung von Lehrkräften zu einer optimalen Unterstüt-

zung beim Lernprozess und einer geeigneten Auswahl an herausfordernden Aufgaben im Unterricht führt, was letztendlich die Schüleroutcomes positiv beeinflussen kann (Behrmann & Souvignier, 2015; Dubberke et al., 2008).

Neben Überzeugungen werden auch die Werthaltungen (Baumert & Kunter, 2006) und Ziele (Schiefele & Schaffner, 2015) von Lehrkräften als bedeutsam für die professionelle Kompetenz von Lehrpersonen angesehen. Da diese in der vorliegenden Arbeit jedoch eine untergeordnete Rolle spielen, wird hier nicht näher auf diese Aspekte eingegangen.

Die empirische Befundlage deutet daher darauf hin, dass der professionelle Kompetenzaspekt „Überzeugungen/Werthaltungen/Ziele“ als bedeutsam für die Unterrichtsgestaltung sowie die Schülerleistungsentwicklung angesehen werden kann. Insbesondere eine konstruktivistische Überzeugung bezüglich Lehr-Lernprozessen scheint positiv mit der Schülerleistungsentwicklung assoziiert zu sein und kann daher als wünschenswerte Eigenschaft von (angehenden) Lehrkräften angesehen werden.

4.2 Motivationale Orientierung

Die zweite Teilkompetenz, die im Modell der professionellen Kompetenz von Lehrkräften genannt wird, ist die motivationale Orientierung von Lehrkräften. Damit sind individuelle Unterschiede bezüglich Zielen, Präferenzen und Motiven gemeint, die – in Abhängigkeit von weiteren Persönlichkeitsmerkmalen sowie dem situationalen Kontext – bedingen, mit welcher Intensität, Qualität oder Dauer Verhaltensweisen gezeigt werden (Kunter, 2011, S. 259). Bezüglich des Lehrberufs können verschiedene Motivationsbereiche unterschieden werden, von denen anzunehmen ist, dass sie das berufliche Handeln und damit auch die Unterrichtsqualität mitbestimmen:

Die Berufswahlmotivationsforschung beschäftigt sich mit den subjektiven Gründen und Motiven, warum sich Schulabgänger dazu entschließen, ein Lehramtsstudium aufzunehmen. Aus Forschungsarbeiten ist bekannt, dass es vorrangig intrinsische Motivationsaspekte wie Interesse und Freude an der Zusammenarbeit mit Kindern und Jugendlichen sind, die die Wahl eines Lehramtsstudiums bedingen (König, Rothland, Darge, Lünemann & Tachtsoglou, 2013). Bezüglich der Be-

rufswahlmotivation sind zudem Unterschiede zwischen verschiedenen Lehramtsstudiengängen zu beobachten: Studierende des Gymnasiallehramts geben beispielsweise häufiger an, ein hohes fachliches und wissenschaftliches Interesse zu haben, während Studierende des Grund-, Haupt-, Real- und Sonderschullehramts ein höheres soziales Interesse berichten (Neugebauer, 2013). Inwiefern sich bestimmte motivationale Konstellationen zu Beginn des Studiums auf die spätere Motivation im Beruf auswirken, kann aufgrund eines Mangels an empirischen Evidenzen bislang nur gemutmaßt werden. Daher stellt die Durchführung von entsprechenden Längsschnittstudien ein wichtiges Forschungsdesiderat dar.

Fest steht allerdings, dass sich motivationale Aspekte kurz- und langfristig positiv auf das Erreichen akademischer und beruflicher Ziele auswirken (Harackiewicz, Barron, Tauer & Elliot, 2002). Zuletzt haben Pohlmann und Möller (2010) einen standardisierten Fragebogen zur Erfassung von Berufswahlmotiven von Lehramtsstudierenden vorgelegt.

Ein weiterer Bereich der Lehrermotivationsforschung befasst sich mit der Motivation im Unterricht. Es liegen hierzu gesicherte Befunde vor, dass die Begeisterung, die Lehrkräfte für ihren Beruf empfinden (sog. *teacher enthusiasm*), positiv mit der Begeisterung und dem Interesse der Schüler für dieses Fach zusammenhängen (Frenzel, Goetz, Lüdtke, Pekrun & Sutton, 2009; Keller, Goetz, Becker, Morger & Hensley, 2014; Kim & Schallert, 2014). In der COACTIV-Studie liegt der Fokus bezüglich motivationaler Orientierung daher auf der Erfassung von *teacher enthusiasm* (Kunter, 2011; Long & Hoy, 2006).

Die Autoren der COACTIV-Studie nennen bezüglich motivationaler Orientierungen ferner die Selbstwirksamkeit und das berufsbezogene Selbstkonzept. Bezüglich der Selbstwirksamkeit liegen empirische Evidenzen vor, dass diese einen Einfluss auf das Instruktionsverhalten von Lehrkräften und somit auch auf die Unterrichtsqualität hat (Holzberger, Philipp & Kunter, 2014). Ferner steht eine hohe Selbstwirksamkeit in einem positiven Zusammenhang mit dem Einsatz von innovativen Lehrmethoden im Unterricht (Tschannen-Moran, Woolfolk Hoy & Hoy, 1998). In weiteren Studien konnte gezeigt werden, dass Lehrkräfte mit einer hohen Selbstwirksamkeitserwartung weniger stressanfällig sind (Abele & Candova, 2007) und besser mit Problemen und Anforderungen umgehen können, die während des Unterrichts auftreten (Überblick: Schwarzer & Warner, 2014, S. 666).

Bezüglich des Zusammenhangs zwischen Lehrerselbstwirksamkeit und Schüleroutcomes sind durchgängig positive Zusammenhänge zu beobachten, beispielsweise zur Motivation der Schüler (Midgley, Feldlaufer & Eccles, 1989) und zur Schülerleistung (Ross, 1992).

Abele (2011) hat den Einfluss des berufsbezogenen Selbstkonzeptes von Mathematiklehrkräften auf den objektiven und subjektiven Berufserfolg über einen Zeitraum von zehn Jahren untersucht. Der objektive Berufserfolg wurde in dieser Studie über den Beschäftigungsumfang operationalisiert, der subjektive Berufserfolg über die Arbeitszufriedenheit und die subjektive Leistungseinschätzung. Zur Erfassung des berufsbezogenen Selbstkonzeptes wurden die Skalen *Agency* („selbstsicher“, „aktiv“, „unabhängig“) und *Communion* („verständnisvoll“, „freundlich“, „hilfsbereit“) eingesetzt (Runge, Frey, Gollwitzer, Helmreich & Spence, 1981). Es konnte gezeigt werden, dass die Skala zu Agency sowohl den objektiven als auch den subjektiven Berufserfolg über einen Zeitraum von zehn Jahren hinweg prognostizieren kann. Die Skala Communion kann ebenfalls Arbeitszufriedenheit über den Zeitraum von zehn Jahren vorhersagen, nicht jedoch den objektiven Berufserfolg.

Zur Erfassung des berufsbezogenen Selbstkonzeptes von Lehramtsstudierenden haben Retelsdorf, Bauer, Gebauer, Kauper und Möller (2014b) einen standardisierten Fragebogen vorgelegt, der sechs inhaltliche Dimensionen differenziert (u. a. Selbstkonzept Fach, Selbstkonzept Erziehung etc.). Es konnten hier Zusammenhänge zwischen den Dimensionen des berufsbezogenen Selbstkonzeptes und der Berufswahlmotivation (Pohlmann & Möller, 2010) der Studierenden gezeigt werden. Weiterhin zeigten sich geringe korrelative Zusammenhänge zwischen der Note im ersten [und zweiten] Unterrichtsfach und dem Selbstkonzept Fach ($r = -0.29$ [$r = -0.23$], alle $ps \leq 0.001$) sowie dem Selbstkonzept Erziehung ($r = -0.16$ [$r = -0.11$], alle $ps \leq 0.001$). Des Weiteren konnte gezeigt werden, dass ein positives berufsbezogenes Selbstkonzept von Lehrkräften dazu führt, dass diese ihre Schüler stärker motivieren, sich vertieft mit den Lerninhalten auseinanderzusetzen, als Lehrkräfte mit einem negativen Selbstkonzept (Yeung, Craven & Kaur, 2014).

Resümierend kann festgestellt werden, dass kaum längsschnittliche Analysen zum Einfluss von motivationalen Faktoren auf die berufliche Entwicklung von (ange-

henden) Lehrkräften vorliegen. Allerdings verdichten sich die bisherigen empirischen Evidenzen darauf, dass ein hohes Ausmaß an *teacher enthusiasm*, eine hohe Selbstwirksamkeitserwartung sowie ein positives berufsbezogenes Selbstkonzept als günstige Entwicklungsfaktoren für die Professionalisierung im Lehrberuf angesehen werden können.

4.3 Selbstregulation

Die dritte Teilkompetenz des Modells der professionellen Kompetenz von Lehrkräften bezieht sich auf die selbstregulativen Fähigkeiten von Lehrkräften (Klusmann, 2011).

„Professionelle Selbstregulationsfähigkeit bedeutet, dass Berufsenagement und überfachliche Problemlöseressourcen (Distanzierungsfähigkeit, offensive Problembewältigung, geringe Resignationstendenz bei Misserfolg und innere Gelassenheit), unterstützt durch mehr oder weniger ausgeprägte berufsbegleitende Emotionen des Lebensgefühls als ‚balanced commitment‘ (Hallsten, 1993) realisiert werden, statt in berufsbezogene Frustration und Entfremdung zu münden“ (Albisser, Kirchhoff & Albisser, 2009, S. 265; zitiert nach Hallsten, 1993).

Im Lehrberuf spielt vor allem der Umgang mit Belastungen und Stress eine bedeutsame Rolle, da Lehrer zu einer der Berufsgruppen zählen, die besonders unter Belastungen leiden (Abele & Candova, 2007). Als besonders belastend wird das Verhalten schwieriger Schüler, die Anzahl an zu unterrichtenden Stunden sowie die Klassenstärke empfunden (Schaarschmidt, 2005). Die subjektiv erlebte Belastung durch den Beruf kann die Qualität des Unterrichts negativ beeinflussen (Vandenberghe, 2006). Ein Gefühl der dauerhaften Erschöpfung und Anspannung auf Seiten der Lehrkräfte hat einen direkten negativen Einfluss auf die Noten der Schüler, das Abschneiden der Schüler in standardisierten Leistungstests, die Schulzufriedenheit der Schüler sowie die wahrgenommene Unterstützung durch den Lehrer (Arens & Morin, 2016). Da die Selbstregulationsfähigkeit sowohl auf das Wohlbefinden der Lehrkräfte als auch auf das Unterrichtshandeln positive Effekte hat, ist ein ausreichendes Maß an Fähigkeiten zur Selbstregulation von enormer Bedeutung für das erfolgreiche Ausüben des Lehrberufs (Roloff Henoch, Klusmann, Lüdtke & Trautwein, 2015). Hieraus kann geschlussfolgert werden,

dass neben bestimmten Überzeugungen und motivationalen Orientierungen auch die Fähigkeit zur Selbstregulation einen fundamentalen Beitrag zum erfolgreichen Lehrerhandeln leistet.

Das Hauptaugenmerk des Modells der professionellen Kompetenz von Lehrkräften liegt jedoch auf der vierten Teilkompetenz, dem professionellen Wissen von Lehrkräften. Da das Konzept des professionellen Wissens von Lehrkräften zentral für diese Dissertationsschrift ist, wird diese Teilkompetenz in dem nachfolgenden Kapitel ausführlich dargestellt.

5 Professionelles Wissen von Lehrkräften

Der Teilkompetenz *Professionelles Wissen* wird innerhalb des Rahmenmodells der professionellen Kompetenz von Lehrkräften eine besonders bedeutsame Rolle zugewiesen (Kunina-Habenicht et al., 2013); bezogen auf Mathematiklehrkräfte scheint dieses nämlich einer der wichtigsten Einflussfaktoren auf die Unterrichtsqualität zu sein (Blömeke & Delaney, 2012). In der Literatur besteht ein Konsens darüber, dass das professionelle Wissen von Lehrkräften nicht als einheitliches Konstrukt verstanden werden kann (Baumert & Kunter, 2006): Es werden einerseits voneinander unterscheidbare inhaltliche Bereiche von professionellem Lehrerwissen angenommen. Andererseits wird die Annahme vertreten, dass unterschiedliche Wissensbereiche des professionellen Lehrerwissens (möglicherweise) unterschiedlich mental repräsentiert werden. Im wissenschaftlichen Diskurs werden also Annahmen diskutiert, die eine inhaltliche bzw. eine formale Kategorisierung von professionellem Lehrerwissen vorschlagen. Eine einheitliche Meinung über die Kategorisierung oder die Topologie des professionellen Wissens von Lehrkräften gibt es bisher u.a. aufgrund mangelnder empirischer Evidenzen nicht (vgl. Baumert & Kunter, 2006).

In diesem Kapitel werden prominente Annahmen zur inhaltlichen, formalen sowie zur inhaltlich-formalen Kategorisierung des professionellen Wissens von Lehrkräften vorgestellt. Des Weiteren soll erläutert werden, inwiefern professionelle Wissensinhalte von Lehrkräften deren professionelles Handeln im Unterricht beeinflussen; denn nur wenn professionelle Wissensinhalte auch im professionellen Handeln praktisch umgesetzt werden, kann von einer Wirkung des professionellen Wissens auf die Unterrichtsqualität und die Schülerleistung gesprochen werden.

5.1 Inhaltliche Kategorisierung des professionellen Wissens von Lehrkräften

Bezüglich der inhaltlichen Kategorisierung des professionellen Wissens von Lehrkräften hat sich in der Literatur die häufig rezipierte theoretische Wissenstaxonomie von Shulman (1986, 1987) durchgesetzt. Shulman postuliert drei professionelle Wissensdomänen von Lehrkräften – das pädagogische Wissen (pedagogi-

cal knowledge, PK³), das Fachwissen (content knowledge, CK) und das fachdidaktische Wissen (pedagogical content knowledge, PCK)⁴ – die als äußerst bedeutsam für das erfolgreiche Lehrerhandeln gelten. Diese drei Wissensdomänen bilden den Kern des Professionswissens von Lehrkräften und „es besteht kein Zweifel darin, dass allen dreien eine zentrale Bedeutung bei den professionellen Aufgaben von Lehrerinnen und Lehrern zukommt“ (Krauss et al., 2008).

Unter PK versteht Shulman (1987) allgemeine pädagogische Fähigkeiten und Fertigkeiten von Lehrkräften, wie Classroom Management. PK wird als fächerübergreifender allgemeiner Wissensinhalt angelegt.

Unter CK versteht Shulman (1987) die fachlichen Wissensinhalte, die sich Lehrkräfte über die Zeit ihrer Ausbildung hinweg aneignen. Das Vorhandensein eines ausreichenden Maßes an CK gilt als notwendige, aber nicht hinreichende Voraussetzung für die Entwicklung an PCK (Baumert & Kunter, 2006; Özden, 2008; Shulman, 1987).

Unter PCK versteht Shulman (1987) „that special amalgam of content and pedagogy that is uniquely the province of teachers, their own special form of professional understanding“ (Shulman, 1987, S. 8). PCK – als Verschmelzung aus pädagogisch-psychologischem- und Fachwissen – gilt als *der* spezifische Wissensinhalt, der für das professionelle und effiziente Handeln von Lehrkräften so bedeutsam ist (Voss et al., 2015). Des Weiteren soll der Wissensbestand PCK Lehrkräfte von Fachwissenschaftlern (z. B. in den Naturwissenschaften) oder von pädagogischen Fachkräften (z. B. in der Erziehungsberatung) differenzieren (Cochran, DeRuiter & King, 1993).

Folgerichtig ist für das erfolgreiche professionelle Handeln im Unterricht zentral, dass Lehrkräfte ihr domänenspezifisches Fachwissen (CK) in einer schülerange-

³ Die englischen Bezeichnungen PPK, CK und PCK werden in dieser Arbeit als Synonyme der deutschen Bezeichnungen pädagogisch-psychologisches Wissen, Fachwissen und fachdidaktisches Wissen verstanden. Aus Gründen der Einheitlichkeit und der besseren Lesbarkeit werden die, in den wissenschaftlichen Diskurs eingebetteten, englischen Abkürzungen verwendet.

⁴ In seiner Originalarbeit nennt Shulman (1987, S. 8) noch vier weitere professionelle Wissensdomänen von Lehrkräften („curriculum knowledge“, „knowledge of learners and their characteristics“, „knowledge of educational contexts“ und „knowledge of educational ends, purposes and values“. Diese werden im Folgenden nicht als professionelle Wissensinhalte im engeren Sinne angesehen und daher nicht weiter thematisiert.

messenen Art und Weise (PK) unter Berücksichtigung typischer – im Einzelfall auch individueller – Schülerkognitionen vermitteln können (PCK).

Das Konzept PCK wurde und wird international vielfach diskutiert und untersucht (Borowski & Riese, 2010). Es wird zwar angenommen, dass einerseits fachliche und andererseits pädagogische Kenntnisse und Fähigkeiten die Grundlage für die Entwicklung von PCK darstellen (van Dijk & Kattmann, 2010), trotzdem werden in der Literatur abgrenzbare professionelle Wissensinhalte angenommen (Baumert et al., 2010). Aufgrund der aktuellen Evidenzlage in der empirischen Bildungsforschung kann davon ausgegangen werden, dass die oben dargestellten professionellen Wissensinhalte – CK, PK und PCK – in bedeutsamer Weise zum Output des Bildungssystems beitragen (Baumert & Kunter, 2013; Hill, Rowan & Ball, 2005).

Es liegt bereits eine gesicherte Befundlage zum positiven Einfluss von PCK auf die Unterrichtsqualität im Fach Mathematik (Hill et al., 2005; Hill et al., 2008; Kunter, Baumert et al., 2011) und den Naturwissenschaften (u. a. Olgetree, 2008; Olszewski, 2010) vor; so konnte beispielsweise gezeigt werden, dass PCK im Fach Mathematik einen Einfluss auf die Unterrichtsqualität und in Folge dessen auf den Lernerfolg von Schülern hat (Baumert & Kunter, 2011a). Die Beschreibung und Erfassung der Lehrerkompetenz, insbesondere der professionellen Wissensinhalte wie PCK, ist aufgrund dessen ein zentrales Anliegen der empirischen Bildungsforschung der letzten fünfzehn Jahre (Helmke, 2012).

Das professionelle Wissen von angehenden Lehrkräften wird in bedeutsamem Maße durch die Ausbildung an der Hochschule bzw. der Universität bestimmt (Kleickmann et al., 2013). Daher kann im Sinne einer Qualitätssicherung bzw. einer Qualitätssteigerung in der Ausbildung von Lehrkräften insbesondere an dieser Stelle Einfluss genommen werden.

Shulman (1986; 1987) legt daher eine theoretische Wissenstaxonomie des professionellen Wissens von Lehrkräften vor, welche die gesamte Forschungslandschaft im Bereich der Lehrerbildungsforschung beflügelte. Insbesondere die Wissensdomäne PCK wird von empirischen Forschungsarbeiten zur Lehrerbildung fokussiert (Depaepe, Verschaffel & Kelchtermans, 2013).

Bei Shulman (1986; 1987) finden sich allerdings keine Vermutungen darüber, wie diese inhaltlichen Bereiche des professionellen Wissens von Lehrkräften mental

repräsentiert werden. Dieser Frage wandten sich nachfolgende Autoren zu, die Theorien zur formalen Kategorisierung von professionellem Lehrerwissen vorlegten.

5.2 Formale Kategorisierung des professionellen Wissens von Lehrkräften

Bromme (1992, 2014) hat sich in seiner wegweisenden Arbeit „Der Lehrer als Experte“ unter dem Gesichtspunkt der psychologischen Lehrerkognitions- und Expertenforschung ausführlich mit dem professionellen Wissen von Lehrpersonen beschäftigt. Er stellt in seiner Arbeit einen Bezug zwischen der Entwicklung des professionellen Wissens von Lehrkräften und der (bereits sehr elaborierten) Forschung zum Expertenwissen innerhalb anderer Berufsgruppen her.

Ein viel beachteter Ansatz innerhalb der kognitionspsychologischen Forschung zum Expertenwissen ist, eine Unterscheidbarkeit von sog. expliziten und impliziten Wissensinhalten anzunehmen (DeJong & Ferguson, 1996; Polanyi, 1985). Explizite Wissensinhalte werden als Wissen verstanden, das verbalisierbar ist und auf Fakten, Regeln und dokumentierten Erfahrungen basiert. Explizite Wissensinhalte sind dem Bewusstsein direkt zugänglich und daher eindeutig verbal reproduzierbar und ohne weiteres klar zu kommunizieren. „Explizites Wissen ist [...] methodisch, systematisch und liegt in artikulierter Form vor. Es ist außerhalb der Köpfe einzelner Personen in Medien gespeichert und kann u. a. mit Mitteln der Informations- und Kommunikationstechnologie aufgenommen, übertragen und gespeichert werden.“ (North, Reinhardt & Sieber-Suter, 2012).

Anders verhält es sich mit impliziten Wissensinhalten: Darunter wird Wissen verstanden, das dem Handeln *unbewusst* zugrunde liegt. Dem impliziten Wissen wird daher eine spezifische Bewusstseinsqualität zugeschrieben. Außerdem „ist implizites Wissen sprachlich-explizit nicht oder nur annäherungsweise zu fassen“ (Bräuer, 2010, S. 323). Polanyi (1985) bezieht die Konzeptualisierung von implizitem Wissen auf die Tatsache „dass wir mehr wissen, als wir zu sagen wissen.“ (S. 14).

Implizites Wissen ist vorrangig an konkrete Erfahrungen und Erlebnisse von einzelnen Personen geknüpft. Es basiert auf bruchstückhaften, subjektiven Erfah-

rungswerten und Erlebnissen, welche unzureichend verbalisiert im Erleben einer Person gespeichert und mit deren Werten und Überzeugungen verwoben sind. Das implizite Wissen liegt im Gegensatz zum expliziten Wissen personengebunden vor; aufgrund dessen ist es äußerst schwierig, implizites Wissen systematisch zu erfassen und zu untersuchen.

Basierend auf den theoretischen Überlegungen von Polanyi (1985) unterscheidet Bromme (2014) bezogen auf das professionelle Wissen von Lehrkräften daher explizite und implizite Wissensbereiche. So basiert das professionelle Wissen von Lehrkräften laut Bromme auf „bewusst gelernte[n] Fakten, Theorien und Regeln [explizites Wissen], sowie [den] Erfahrungen und Einstellungen [implizites Wissen] des Lehrers“ (Bromme, 2014, S. 10).

5.3 Inhaltlich-formale Kategorisierung des professionellen Wissens von Lehrkräften

Die beiden vorherigen Kapitel beschreiben Ansätze zur inhaltlichen (siehe 5.1) sowie formalen (siehe 5.2) Kategorisierung von professionellem Lehrerwissen. Des Weiteren gibt es in der Literatur Bestrebungen, durch eine Verknüpfung von inhaltlicher und formaler Kategorisierung des professionellen Lehrerwissens, eine Synthese aus beiden Ansätzen zu bilden. So unterscheidet Fenstermacher (1994) bezüglich des professionellen Wissens von Lehrkräften zwei Wissensbereiche: Das theoretisch-formale Wissen (*formal knowledge*) und das praktische Wissen und Können (*practical knowledge*).

Das theoretisch-formale Wissen meint in erster Linie das fachliche Wissen von Lehrpersonen; es umfasst jedoch auch Teile des fachdidaktischen und allgemeinen pädagogischen Wissens von Lehrpersonen.

Unter praktischem Wissen und Können versteht Fenstermacher (1994) erfahrungsbasiertes Wissen von Lehrkräften, das sich auf konkrete Problemstellungen und spezifische Kontexte, beispielsweise bestimmte Unterrichtssituationen, bezieht. Praktisches Wissen und Können bildet die Grundlage für promptes professionelles Handeln im Unterricht und liegt in der Regel als impliziter Wissensinhalt vor.

Fenstermacher (1994) verbindet in der Unterscheidung der beiden Wissensbereiche – theoretisch-formales Wissen und praktisches Wissen und Können – eine inhaltliche sowie eine formale Kategorisierung von professionellem Lehrerwissen. Das theoretisch-formale Wissen, das vorrangig aus fachlichem Wissen besteht, das die (angehenden) Lehrkräfte durch ihre Ausbildung an der Hochschule erwerben, ist als ein expliziter Wissensinhalt zu verstehen, der eindeutig verbalisierbar und klar zu kommunizieren ist. Fachliches Wissen – laut Shulmans Taxonomie CK – sei daher ein expliziter Wissensinhalt. Das praktische Wissen und Können wird als erfahrungsabhängig verstanden und ist damit als ein personengebundener, impliziter Wissensinhalt zu verstehen.

Wenn man die vorgeschlagenen Überlegungen zu formalen sowie formal-inhaltlichen Kategorisierungen von professionellem Lehrerwissen resümiert und mit der weitläufig akzeptierten inhaltlichen Kategorisierung des professionellen Lehrerwissens von Shulman (1986; 1987) in Bezug setzt, dann können bezüglich der mentalen Repräsentation der einzelnen inhaltlichen Wissensdomänen folgende Annahmen getroffen werden:

CK kann als ein expliziter Wissensinhalt von Lehrkräften verstanden werden. CK bezieht sich auf domänenspezifisches Fach- und Faktenwissen, das im Rahmen der theoretischen Ausbildung an der Hochschule in der ersten Phase der Lehrerbildung vermittelt und von den Lehramtsstudierenden erworben wird.

Sowohl PK als auch PCK beziehen sich *nicht* ausschließlich auf theoretisch erworbenes Fach- und Faktenwissen und werden z. T. durch eigene Erfahrungen erworben (zu PCK: Lee & Luft, 2008). Plausibel ist daher anzunehmen, dass die Wissensdomänen PK und PCK teilweise als implizite Wissensinhalte repräsentiert werden (Alonzo & Kim, 2015). Im wissenschaftlichen Diskurs gibt es bislang allerdings kaum eine breite Diskussion darüber, in welchen Verhältnissen PK und PCK aus impliziten und expliziten Wissensanteilen bestehen. Des Weiteren ist vollkommen offen, wie sich die impliziten und expliziten Anteile von PK und PCK gegenseitig bedingen. Neuere Ansätze, die sich mit der Verknüpfung von impliziten und expliziten Wissensbeständen beschäftigen, gehen dabei nicht von zwei unabhängigen Wissenssystemen aus, sondern postulieren einen engen Zusammenhang von impliziten und expliziten Wissensbeständen (Berry, Shanks & Henson, 2008).

Hieraus folgt, dass es für die systematische empirische Untersuchung von professionellem Lehrerwissen von zentraler Bedeutung ist, die postulierten mentalen Repräsentationsformen der professionellen Wissensinhalte von Lehrkräften zu berücksichtigen. Insbesondere bei der Erfassung von PCK und PK sollte beachtet werden, dass diese inhaltlichen Wissensdomänen sowohl als implizite als auch als explizite Wissensbestände vorliegen (können).

5.4 Zusammenhang zwischen professionellem Wissen und professionellem Handeln

In die Diskussion um den Erwerb professioneller Wissensinhalte im Lehramtsstudium müssen selbstverständlich auch Überlegungen eingebracht werden, wie sich (und ob überhaupt) die im Studium erworbenen professionellen Wissensinhalte im späteren professionellen Handeln im Unterricht niederschlagen.

Studien, die sich mit dieser Fragestellung beschäftigt haben, zeigen ernüchternde Ergebnisse: Lehrkräfte greifen nach eigenen Angaben im unterrichtlichen Handeln höchst selten auf im Studium erworbenes Wissen zurück und zwar auch dann, wenn sie sich in Unterrichtssituationen ohne besonderen Handlungsdruck befinden. Die große Mehrheit von über 500 befragten Lehrkräften (81.7 %) gibt an, dass sie sich bei pädagogischen Entscheidungen ohne Zeitdruck an den eigenen beruflichen Erfahrungen orientieren; nur eine äußerst geringe Zahl (6.4 %) nennt bei dieser Frage Inhalte, die im Studium vermittelt wurden (Terhart, 1994).

Welche Gründe kann es dafür geben, dass professionelle Wissensinhalte, die im Studium vermittelt und erworben werden, nicht unmittelbar in professionelles Handeln „übersetzt“ werden können? Eben diese Verzahnung zwischen theoretischer Ausbildung und praktischem Handeln ist außerordentlich bedeutsam für das effiziente und professionelle Handeln im Unterricht und muss daher als *das* Ziel der Lehramtsausbildung gelten.

Diesbezüglich wurde angeregt, dass Lehramtsstudierende neben den bloßen Wissensbeständen – explizites Wissen – über ihre Ausbildung hinweg eine Kontextualisierungskompetenz – implizites Wissen – erwerben müssen (Blömeke, König, Suhl, Hoth & Döhrmann, 2015; Neuweg, 2005; Patry, 2000). Das bedeutet, dass die Studierenden erlernen sollten, die im Studium erworbenen Wissensbestände

kontextbasiert einzusetzen. Ein Beispiel hierfür wäre, dass die Studierenden erlernen zu unterscheiden, wann es sich um die Ausnahme von einer gewissen Verhaltensregel handelt. Unter welchen Bedingungen ist es beispielsweise sinnvoll, eine Lernaufgabe abubrechen? Wie sollte man in bestimmten Situationen mit Unterrichtsstörungen umgehen?

Neben der Fähigkeit professionelles Wissen kontextbasiert einzusetzen ist auch die Prozeduralisierung dieses Wissens eine notwendige Bedingung für erfolgreiches professionelles Handeln (Neuweg, 2005). Das bedeutet, dass das professionelle Handeln im Unterricht auch stetig (ein)geübt werden muss, was allerdings in der Hochschuldidaktik kaum eine Rolle spielt (Bollonow, 1991; Neuweg, 2005).

Allerdings kann auch nicht davon ausgegangen werden, dass praktische Erfahrungen automatisch zu einer optimalen Entwicklung der professionellen Kompetenz führen. So wird beispielsweise berichtet, dass es keinen „signifikanten Zusammenhang zwischen der Dauer der Berufstätigkeit von Lehrern, dem Niveau ihres Expertenwissens und ihrem Unterrichtserfolg“ gibt (Weinert & Helmke, 1996, S. 232). Ein positiver Zusammenhang zwischen der Dauer der Berufstätigkeit und dem Ausmaß an professioneller Kompetenz wurde für andere Expertisedomänen wie Medizin, Schach und Physik hingegen gezeigt (Weinert & Helmke, 1996). Eine Ursache für den fehlenden Zusammenhang zwischen der Dauer der Berufstätigkeit und dem Ausmaß an professioneller Kompetenz bei Lehrkräften könnte sein, dass Schulklassen nicht immer ein zuverlässiges Rückmeldesystem darstellen, um eine langfristige und stetige Kompetenzsteigerung bei Lehrkräften zu erwirken (Weinert & Helmke, 1996).

In der Lehramtsausbildung sowie bei der Gestaltung von Fort- und Weiterbildungsmaßnahmen für (erfahrene) Lehrkräfte könnte daher die Etablierung eines professionellen Austauschs unter Kollegen (bzw. unter Kommilitonen) ein vielversprechender Ansatzpunkt sein, um den Aufbau professioneller Kompetenzen im Lehrerberuf zu begünstigen. Bonsen und Rolff (2006) stellen hierzu das in den angelsächsischen Ländern weit verbreitete Konzept der „Professionellen Lerngemeinschaft für Lehrerinnen und Lehrer“ vor. Dieses Konzept zielt eben darauf ab, den professionellen kollegialen Austausch zwischen den Lehrern innerhalb einer Schule zu fördern.

Neuweg (2014) hat den Zusammenhang zwischen professionellem Wissen und professionellem Handeln im Lehrberuf in seinem Modell „Konzepte des ‚Lehrerwissens‘“ dargestellt (siehe Abbildung 2).

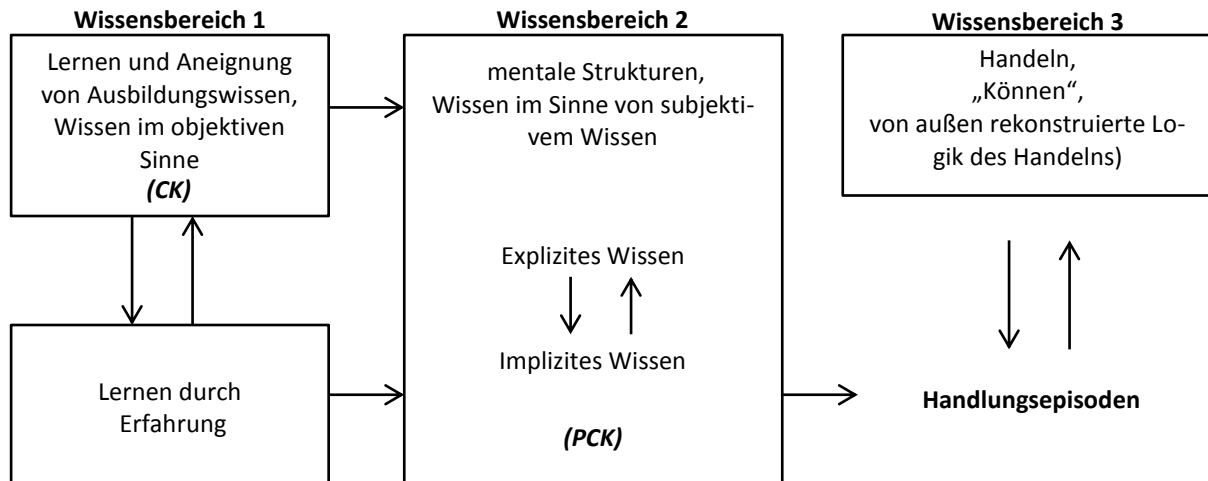


Abbildung 2. Konzepte des „Lehrerwissens“. Eigene und erweiterte Darstellung nach Neuweg (2014, S. 585).

Wissensbereich 1 bezieht sich auf das Ausbildungs- und Faktenwissen, das Lehramtsstudierende im Rahmen ihres Studiums erwerben. Im Sinne von Shulman (1987) kann Wissensbereich 1 daher inhaltlich als die Wissensdomäne CK verstanden werden. Wissensbereich 1 und das Lernen durch Erfahrungen (u. a. durch eigene Lehrerfahrungen im Unterricht) beeinflussen sich in diesem Modell gegenseitig.

Wissensbereich 1 und das Lernen durch Erfahrung bedingen den Zuwachs von Wissensbereich 2. Unter Wissensbereich 2 versteht Neuweg (2014) mentale Strukturen von (angehenden) Lehrkräften sowie Wissen im Sinne von subjektivem Wissen, das sich sowohl aus impliziten als auch aus expliziten Anteilen zusammensetzt. Wissensbereich 2 kann in Shulmanschen Begrifflichkeiten als PCK verstanden werden (Shulman, 1987).

Wissensbereich 2 bildet die Grundlage für die Entwicklung von Wissensbereich 3, der sich auf das professionelle Handeln im Unterricht bezieht. Wissensbereich 3 beschreibt Neuweg (2014) als Grundlage für das professionelle Handeln im Un-

terricht, da dieser als dispositionale Erklärung für das Verhalten von Lehrkräften im Unterricht verstanden werden kann.

„Wissen[sbereich] 3 wird häufig zur dispositionalen (nicht kausalen!) ‚Erklärung‘ von Verhalten verwendet. Dass beispielsweise ein Lehrer eine Klasse zu führen ‚weiß‘, bedeutet dann, dass er sie führen kann. Ihm ‚Wissen‘ zuzuschreiben, heißt, ihm eine Verhaltensdisposition zuzuschreiben: Er ‚weiß, wie es geht‘ – in dem Sinne, dass er es demonstrieren kann; nicht notwendigerweise muss er es verbalisieren können“ (Neuweg, 2014, S. 584-585).

Neuweg (2014) weist jedoch auch auf die Bruchstellen innerhalb des Modells und die zum Teil verschwimmenden Grenzen zwischen den drei Wissensbereichen hin. So kann beispielsweise nicht kausal von einer Wirkung von Wissensbereich 2 auf Wissensbereich 3 ausgegangen werden.

Das Modell „Konzepte des ‚Lehrerwissens‘“ macht um ein Weiteres die besondere Bedeutung der Wissensdomäne PCK deutlich: PCK wird sowohl durch CK als auch durch Erfahrungslernen (aus)gebildet und stellt einen eigenen, abgrenzbaren professionellen Wissensbereich mit impliziten und expliziten Anteilen dar. PCK wird darüber hinaus als Schnittstelle bzw. Übertragungsmoment von theoretischem Fach- und Faktenwissen und dem professionellen Handeln im Unterricht konzeptualisiert.

Weiterhin lassen sich basierend auf diesen Überlegungen Implikationen für die Lehramtsausbildung ableiten: Anstatt einer allzu strikten Trennung zwischen theoretischer Ausbildung im Lehramtsstudium und der praktischen Ausbildung im Referendariat, sollte vielmehr die Verzahnung von theoretischem Wissen und praktischem Können forciert werden. Es ist bereits im Studium wichtig, konkrete Unterrichtssituationen zu analysieren und zu lernen, das eigene Handeln zu reflektieren; ebenso ist es im Referendariat wichtig, Rückbezüge auf pädagogisch-psychologische Theorien oder fachdidaktische Prinzipien, die im Studium erworben wurden, herzustellen. Diese Überlegungen haben bereits Einzug in die Ausbildung der Lehramtsstudierenden gehalten, da z. B. an den Pädagogischen Hochschulen in Baden-Württemberg im Rahmen eines, durch Dozierende begleitetes, Semesterpraktikums eben diese Fähigkeiten der Studierenden gefördert werden sollen (Pädagogische Hochschule Heidelberg, 2016).

5.5 Entwicklung des professionellen Wissensinhalts PCK bei (angehenden) Lehrkräften

Die vorherigen Kapitel haben verdeutlicht, dass PCK als ein bedeutsamer professioneller Wissensinhalt von (angehenden) Lehrkräften verstanden werden kann, der Einfluss auf das professionelle Handeln im Unterricht nimmt. Daher ist für die Lehrerbildung von größtem Interesse, wie sich PCK und professionelles Wissen im Allgemeinen über das Lehramtsstudium und das Referendariat hinweg entwickeln und wie diese Entwicklung bestmöglich gefördert werden kann. Hierzu liegen bislang kaum empirische Forschungsarbeiten vor (u. a. Scherf, 2013). Wenige Studien können echte Längsschnittdaten für entsprechende Analysen zugrunde legen, sodass PCK in der Regel querschnittlich oder quasi-längsschnittlich untersucht wurde. Daher können Aussagen über (individuelle) Entwicklungsverläufe von PCK sowie über Effekte von Fördermaßnahmen bisher nicht gesichert getroffen werden.

Im Folgenden werden zunächst Überlegungen zur allgemeinen Entwicklung von PCK vorgestellt. Anschließend werden empirische Ergebnisse zu verschiedenen Einflussfaktoren auf die Entwicklung von PCK präsentiert.

5.5.1 Allgemeine Entwicklung von PCK

Um sich dem Forschungsgegenstand der allgemeinen Entwicklung von PCK zu nähern, greifen einige Autoren auf theoretische Überlegungen und empirische Evidenzen aus der Expertiseforschung innerhalb anderer Berufsgruppen zurück (Kunze, 2004). Hier gibt es zahlreiche Forschungsarbeiten, die gesicherte Befunde darüber vorlegen, dass sich das Wissen von Novizen und Experten unterscheidet (Überblick: Palmer, Stough, Burdinski Jr. & Gonzales, 2005): U. a. konnte festgestellt werden, dass Experten sich hinsichtlich ihres Denkens und Verhaltens qualitativ von Novizen unterscheiden, dass sich Expertenwissen auf eine spezifische Domäne bezieht und durch viele Stunden intensiver Übung (sog. *deliberate practice*) erworben wird und dass Expertenwissen qualitativ anders strukturiert ist, als das Wissen von Novizen.

Auch in Bezug auf den Lehrerberuf kann basierend auf diesen Forschungsergebnissen davon ausgegangen werden, dass sich das professionelle Wissen von Lehramtsstudierenden (Novizen) und erfahrenen Lehrkräften (Experten) (qualitativ)

unterscheidet. Es wird nämlich vermutet, dass Lehramtsstudierende/Referendare im Unterrichtshandeln auf „einzelne Wissensbestände“ zurückgreifen, wohingegen erfahrenen Lehrkräften „ein teilweise routiniertes Wissen, in dem wissenschaftlich basiertes Wissen, eigenes Erfahrungswissen und tradiertes Wissen des Berufsstandes“ verschmolzen sind, zur Verfügung steht (Kunze, 2004, S. 52).

Dreyfus (2004) hat ein Stufenmodell der Entwicklung vom Novizen zum Experten vorgelegt. Für die Entwicklung vom Novizen zum Experten werden hierin fünf konsekutive Stufen postuliert. Bezogen auf das Beispiel „Lehrerberuf“ wird in diesem Modell angenommen, dass alle angehenden Lehrkräfte im Rahmen des Professionalisierungsprozesses im Lehramtsstudium und Referendariat dieselben fünf Stufen durchlaufen (siehe auch: Berliner, 1988).

Professionalisierungsprozess	Fähigkeitsstufe	Entscheidungskomponente	Professionelle Perspektive	Entscheidungsgrundlage	Engagement/Einsatz
	Expert <i>Erfahrene Lehrkraft</i>	Ohne Kontextbezug oder situational	Erfahren	Intuitiv	Involviert
	Proficient <i>Lehrkraft</i>	Ohne Kontextbezug oder situational	Erfahren	Analytisch	Teilweise losgelöst, teilweise involviert
	Competent <i>Berufseinsteiger nach dem Referendariat</i>	Ohne Kontextbezug oder situational	Zunehmend erfahren	Analytisch	Teilweise losgelöst, teilweise involviert
	Advanced Beginner <i>Praktikanten/Referendare</i>	Ohne Kontextbezug oder situational	Keine	Analytisch	Losgelöst
	Novice <i>Lehramtsstudierende</i>	Ohne Kontextbezug	Keine	Analytisch	Losgelöst

Abbildung 3. Modell „Five Stages of Skill Acquisition“. Eigene und erweiterte Darstellung nach Dreyfus (2004, S. 181).

„Entscheidungskomponente“ meint die Aspekte von professionellen Situationen, die von der handelnden Person wahrgenommen werden (können). Dies können sowohl allgemeine als auch situationsspezifische Charakteristika sein. „Professionelle Perspektive“ meint das Fokussieren von relevanten Situationscharakteristika für das professionelle Handeln. „Entscheidungsgrundlage“ für professionelles Handeln können analytische Schlussfolgerungen sein oder eine intuitive Entscheidung, die auf Erfahrungen beruht. „Engagement/Einsatz“ beschreibt in welchem Ausmaß die handelnde Person sich in professionelle Aufgaben und Entscheidungen eingebunden fühlt.

Der Novize, im vorliegenden Fall der Lehramtsstudierende, urteilt in einer professionellen Situation basierend auf erlerntem Faktenwissen bzw. Verhaltensregeln, ohne relevante Kontextmerkmale berücksichtigen zu können. Dreyfus (2004) schreibt hierzu: „The student needs not only the facts but also an understanding of the context in which that information makes sense.” (Dreyfus, 2004, S. 177). Er ist außerdem nicht in der Lage die Komplexität der Situation vollständig zu erfassen und kann auch keine professionelle Perspektive auf das Geschehen einnehmen. Der Novize kommt durch analytisches Schlussfolgern zu einer Lösung für die berufsbezogene Problemstellung und fühlt sich auch wenig eingebunden in das eigene professionelle Handeln.

Der Experte, im vorliegenden Fall der erfahrene Lehrer, kann beim Urteilen in professionellen Situationen auf relevante situationale Merkmale zurückgreifen: Er überblickt das Unterrichtsgeschehen oder die Beurteilungssituation in ihrer gesamten Komplexität. Er kann daher eine professionelle Perspektive auf das Geschehen einnehmen und trifft die Handlungsentscheidung intuitiv (Anmerkung: also auf Grundlage von Wissensinhalten, die z. T. implizit vorliegen, siehe Kapitel 5.2). Des Weiteren fühlt sich der Experte stark eingebunden in sein professionelles Handeln.

Zwischen den Entwicklungsstufen des Novizen und des Experten werden drei Zwischenstufen angenommen. Von Stufe zu Stufe finden Veränderungen hinsichtlich der Entscheidungsfindung, der Perspektive auf professionelle Situationen, des Urteilsprozesses und der Einbindung in diese beruflichen Entscheidungen statt. Fokussiert man die professionelle Entwicklung von Beginn bis zum Ende des Lehramtsstudiums ist allenfalls mit dem Erreichen der zweiten („Advanced Beginner“) oder dritten Stufe („Competent“) zu rechnen. Durch die hauptsächlich theoretische Beschäftigung mit fachlichen und fachdidaktischen Inhalten zu Beginn des Studiums kann davon ausgegangen werden, dass die Studierenden langsam und kontinuierlich ansteigend erlernen (z. B. durch den Besuch anwendungsorientierter Seminare), auch Kontextinformationen mit in ihren Entscheidungsprozess einfließen zu lassen. Durch eigene Praxiserfahrung im Unterrichten (beispielsweise im Integrierten Semesterpraktikum (ISP)) kann außerdem eine zunehmend professionelle Perspektive auf das Unterrichtsgeschehen eingenommen werden.

Die vierte Stufe („Proficient“) kann vermutlich erst durch umfangreiche praktische Erfahrungen als Lehrkraft im Anschluss an das Referendariat erreicht werden, da sich innerhalb dieser Stufe eine zunehmend erfahrene Perspektive sowie eine zunehmende Eingebundenheit in das eigene berufliche Handeln, beispielsweise auch durch das Übernehmen einer eigenen Klasse, entfaltet.

Die fünfte Stufe („Expert“) kann vermeintlich erst nach einigen Jahren der Berufserfahrung erreicht werden und stellt den Endpunkt der professionellen Entwicklung von Lehrkräften dar.

Aus diesen Überlegungen kann geschlussfolgert werden, dass sich PCK – ebenso wie das professionelle Wissen im Allgemeinen – in einem zwar subjektiv verschieden ablaufenden, aber generell zu charakterisierenden Prozess entwickelt (Kunze, 2004).

5.5.2 Einflussfaktoren auf die Entwicklung von PCK

Neben dem allgemeinen Entwicklungsverlauf von PCK interessiert die Lehrerbildungsforschung vertiefend, welche Ausbildungsbausteine und -abschnitte den Zuwachs von PCK in bedeutsamem Maß begünstigen. Diesbezüglich haben verschiedene Autoren angeregt, dass sich die Entwicklung von PCK aus unterschiedlichen, z. T. unabhängigen, Wissens- und Erfahrungsquellen speist (Grossman, 1990; Haston & Leon-Guerrero, 2008; Kind, 2009; van Driel, Verloop & Vos, 1998). So liegen empirische Evidenzen vor, dass die eigene Lehrererfahrung von Lehramtsstudierenden zum Aufbau von PCK beiträgt (Grossman, 1990; Kind, 2009). Es konnte u. a. gezeigt werden, dass PCK in den ersten Monaten der beruflichen Tätigkeit signifikant ansteigt (Simmons et al., 1999). Eine weitere Quelle für den Aufbau von PCK sind spezifische Veranstaltungen an der Hochschule, die domänenspezifisch fachdidaktische Vorgehensweisen thematisieren (Grossman, 1990; Haston & Leon-Guerrero, 2008). Daneben gelten das Vorhandensein eines ausreichenden Maßes an CK, Erfahrungen aus der eigenen Schulzeit sowie kollegiale Zusammenarbeit als bedeutsam für die Entwicklung von PCK (Evens, Elen & Depaepe, 2015). Bisher liegen jedoch kaum empirische Evidenzen dazu vor, wie spezifische Ausbildungsangebote an der Hochschule zu einem Zuwachs an PCK beitragen können (Evens et al., 2015).

In Hinblick auf die Qualitätssicherung und Optimierungsprozesse in der Lehramtsausbildung sind Kenntnisse über die Entwicklung professioneller Wissensinhalte als wichtiger Teilbereich der professionellen Kompetenz von Lehrkräften außerordentlich bedeutsam. Insbesondere Kenntnisse über die Wirkmechanismen sind hierbei entscheidend: Nur wenn bekannt ist, welche Bausteine der Lehramtsausbildung maßgeblich zu einem Aufbau von PCK beitragen, können diese auch gefördert und intensiviert werden, um die Ausbildungsinhalte und den Ausbildungsverlauf zu verbessern und anzupassen.

5.6 Struktur des professionellen Wissensinhalts PCK

Shulman (1986) konzeptualisiert die professionelle Wissensdomäne PCK strukturell in zwei Wissensfacetten und zwar in (1) das Wissen über fachbezogene Schülerkognitionen und (2) das Wissen über Erklären und Darstellen. Nachfolgende Autoren haben diese Strukturierung teilweise umgestaltet und erweitert.

Grossman (1990) beschreibt vier Facetten, die zentral für PCK seien: (1) Wissen über das Schülerverstehen, (2) Wissen über das Curriculum, (3) Wissen über Instruktionsstrategien und (4) Wissen über das Lehrziel.

Cochran et al. (1993) betonen die dynamische Natur von PCK und führen in ihren Forschungsarbeiten daher den Begriff *pedagogical content knowing* (PCKg) ein. Für PCKg werden die Wissensfacetten (1) Pädagogik, (2) fachspezifische Inhalte, (3) Schülercharakteristika und (4) situationaler Lernkontext unterschieden.

Magnusson, Krajacik und Borko (1999) erweitern die ursprüngliche Kategorisierung von Shulman (1986) um drei zusätzliche Facetten: (3) Wissen und Überzeugungen bezüglich Lehrzielen, (4) curriculares Wissen und (5) Wissen zur Erschließung wissenschaftlicher Texte (englisch: *knowledge of assessment of scientific literacy*).

Ball, Thames und Phelps (2008) legen zur inhaltlichen Strukturierung von PCK bei Mathematiklehrkräften eine einflussreiche Konzeptualisierung vor: Die Autoren unterscheiden *mathematical knowlegde for teaching* (MKT) und *content knowledge for teaching mathematics* (CMKT) (Hill et al., 2005; siehe auch: Hill et al., 2008; Hill, Schilling & Ball, 2004). Diese Kategorisierung unterscheidet sich zu der ursprünglichen Kategorisierung von Shulman (1986) dahingehend,

dass keine inhaltliche Trennung von PCK und CK angenommen wird, sondern dass diese beiden Wissensdomänen miteinander verschmolzen vorliegen.

Depaepe et al. (2013) haben ein systematisches Review zur Konzeptualisierung der Struktur von PCK bei Mathematiklehrkräften vorgestellt. Hierin geben die Autoren eine Übersicht über empirische Studien, die die Struktur von PCK untersucht haben. Insgesamt wurden 60 Studien für das Review ausgewählt (siehe Tabelle 1).

Bezüglich der Struktur von PCK werden in diesem Review acht verschiedene inhaltliche Wissensfacetten unterschieden: (1) Schülerfehlvorstellungen und Schülerschwierigkeiten, (2) Instruktionsstrategien und Repräsentationen, (3) Mathematische Aufgaben und kognitive Anforderungen, (4) Bildungsziele, (5) Curriculum und Medien, (6) Kontextwissen, (7) CK und (8) PK. Für die 60 berücksichtigten Studien wurde überprüft, welche Struktur von PCK durch die jeweiligen Autoren postuliert wurde. Außerdem wurde untersucht, welche inhaltlichen Wissensfacetten von PCK explizit in den Studien genannt und operationalisiert wurden.

Tabelle 1. Konzeptualisierung von PCK (eigene Darstellung nach Depaepe et al., 2013).

N	PCK-Modell	Wissensfacetten PCK							
		1	2	3	4	5	6	7	8
16	Shulman (1986, 1987)	x	x						
10	Shulman (1986, 1987)	Keine Spezifikation							
5	MKT	x	x			x			
4	MKT	Keine Spezifikation							
1	Grossman (1990)	x	x		x	x			
1	Cochran et al. (1993)	x					x	x	x
2	An et al. (2004)	x	x			x		x	
1	Veal & MaKinster (1999)	x	x					x	
2		x	x	x					
4		x	x			x			
2		x	x	x				x	
1		x	x		x				
1		x	x			x		x	
10		Keine Spezifikation							

Anmerkungen. N bezieht sich auf die Anzahl an Studien, die im systematischen Review berücksichtigt wurden.

Zur Nennung der Autoren siehe Depaepe et al. (2013).

„PCK-Modell“ meint, dass die berichtete Studie eine bestimmte Konzeptualisierung von PCK zugrunde legt.

Legende: (1) Schülerfehlvorstellungen und Schülerschwierigkeiten, (2) Instruktionsstrategien und Repräsentationen, (3) Mathematische Aufgaben und kognitive Anforderungen, (4) Bildungsziele, (5) Curriculum und Medien, (6) Kontextwissen, (7) CK, (8) PK.

Der Großteil der berichteten Studien (43 %) legt ausdrücklich Shulmans (1986) Konzeptualisierung von PCK mit den beiden inhaltlichen Wissensfacetten (1)

Wissen über fachbezogene Schülerkognitionen und (2) Wissen über Erklären und Darstellen zugrunde.

Daneben kann beobachtet werden, dass alle Studien (bis auf eine Ausnahme) explizit die beiden von Shulman (1986) eingeführten Wissensfacetten von PCK untersuchen⁵. Ergänzend hierzu wurden in 33 % der Studien weitere inhaltliche Wissensfacetten angenommen und untersucht.

Weiterhin weisen Depaepe et al. (2013) darauf hin, dass die Bezeichnungen der inhaltlichen Wissensfacetten von PCK in den verschiedenen Studien nicht einheitlich verwendet werden. So werden beispielsweise die ursprünglichen PCK-Facetten nach Shulman (1986) bei Ball et al. (2008) als *knowledge of content and students* und *knowledge of content and teaching* bezeichnet.

Für andere Fachbereiche – außerhalb der Mathematik bzw. der Naturwissenschaften – liegen keineswegs derart elaborierte Übersichtsartikel zur Struktur von PCK vor. Dieser Fokus der PCK-Forschung auf den Naturwissenschaften könnte u. a. mit dem hohen gesellschaftlichen Stellenwert in Zusammenhang stehen, den die Naturwissenschaften gegenüber den Geisteswissenschaften genießen (Evens et al., 2015). Des Weiteren sind naturwissenschaftliche gegenüber geisteswissenschaftlichen Disziplinen deutlich häufiger Teil von nationalen und internationalen Schulleistungsuntersuchungen (Grossman & Stodolsky, 1995).

Die Ergebnisse des systematischen Reviews von Depaepe et al. (2013) deuten zusammenfassend darauf hin, dass es einen Konsens zwischen verschiedenen Autoren(-gruppen) gibt, dass insbesondere die zwei ursprünglich von Shulman (1986) konzeptualisierten professionellen Wissensfacetten (1) Wissen über fachbezogene Schülerkognitionen und (2) Wissen über Erklären und Darstellen bedeutsame strukturelle Elemente von PCK sind. Diese beiden inhaltlichen Facetten von PCK können daher als Kernelemente des fachdidaktischen Wissens und Könnens angesehen werden (van Dijk & Kattmann, 2010).

Allerdings sind diese Befunde bisher nur in eingeschränktem Maße auf die geisteswissenschaftlichen Domänen generalisierbar, da hier aufgrund eingeschränkter

⁵ Gilt nur für Studien, in denen die untersuchten Wissensfacetten von PCK explizit genannt wurden.

Forschungsbefunde zur Strukturierung von PCK kaum gesicherte evidenzbasierte Aussagen getroffen werden können.

5.7 Testinstrumente zur Erfassung des professionellen Wissensinhalts PCK

Mit dem steigenden Interesse an der systematischen Beforschung des professionellen Wissensinhalts PCK ist ebenso das Interesse an Testverfahren zur Erfassung professioneller Wissensinhalte von Lehrkräften stetig angestiegen. Hinsichtlich der Entwicklung geeigneter Testverfahren ergeben sich allerdings „grundsätzliche methodische Herausforderungen“ (Kunter & Klusmann, 2010).

In Hinblick auf die Erfassung professioneller Kompetenzen im Allgemeinen ist zunächst zwischen Selbst- und Fremdbeurteilungsverfahren zu unterscheiden. Bei Selbstbeurteilungsverfahren (auch: „subjektive Erfassung“) werden die Probanden direkt nach ihrer Einschätzung der eigenen professionellen Kompetenz gefragt (Frey, 2004). Obwohl diese Selbstberichtsdaten ökonomisch mithilfe von Fragebögen erfasst werden können, wird in der Literatur kritisch angemerkt, dass solche Verfahren wenig valide seien (Oser, Curcio & Düggele, 2007) und darüber hinaus nur sog. „träges Wissen“ erfassen würden. Der Begriff des „trägen Wissens“ (*inert knowledge*) geht auf Whitehead (1929) zurück: Unter trägem Wissen versteht er Wissensinhalte, die zwar in bestimmten Situationen abrufbar sind (z. B. beim Ausfüllen eines Fragebogens), in authentischen Situationen aber nicht angewendet werden können (Lindmeier, 2012). Hierzu soll exemplarisch ein kurzes Beispiel gegeben werden: Bezüglich der Kompetenzerfassung anhand von Selbstberichtsdaten könnten befragte Lehramtsstudierende möglicherweise auf die Frage „*Kennen Sie das didaktische Prinzip...?*“ mit „Ja“ antworten, weil sie dieses Prinzip tatsächlich kennen. Die bloße theoretische Kenntnis eines didaktischen Prinzips geht allerdings nicht automatisch mit der Fähigkeit zur Umsetzung des didaktischen Prinzips in der tatsächlichen Unterrichtssituation einher. Wenn ein Testinstrument vorrangig träges Wissen erfasst, sind die resultierenden Testwerte daher als wenig aussagekräftig für die Einschätzung der Kompetenz anzusehen, die eigentlich abgebildet werden soll.

Bei Fremdbeurteilungsverfahren (auch: „objektive Erfassung“) professioneller Kompetenzen werden standardisierte Testinstrumente genutzt. Hierbei kann wei-

terhin unterschieden werden, ob es sich um eine distale oder proximale Erfassung der professionellen Kompetenz handelt (Alisch, Hermkes & Möbius, 2009). Unter einer distalen Erfassung der professionellen Kompetenz versteht man das Heranziehen von konstruktfernen Indikatoren, beispielsweise der Noten von Lehramtsstudierenden im Studium als Maß für professionelle Kompetenz (Baumert & Kunter, 2006). Unter einer proximalen Erfassung versteht man die Verwendung von konstruktnahen Indikatoren zur Abschätzung der professionellen Kompetenz, beispielsweise ein standardisiertes Testinstrument (Desimone, 2009).

Kunter und Klusmann (2010) haben die Übereinstimmung zwischen den Ergebnissen von Selbst- und Fremdbeurteilungsverfahren zur Erfassung der professionellen Kompetenz von Referendaren mit dem Fach Mathematik untersucht. Dazu wurden CK und PCK sowohl im Selbstbericht als auch mithilfe eines standardisierten Testinstruments erfasst (Kunter, Baumert et al., 2011). Der selbsteingeschätzte Indikator für CK weist einen geringen korrelativen Zusammenhang von $r = 0.20$ ($p \leq 0.05$) zum objektiven Indikator für CK auf. Der selbsteingeschätzte und der objektive Indikator für PCK weisen keinen statistisch bedeutsamen Zusammenhang auf. Die Autoren führen diese Diskrepanz auf die geringe Berufserfahrung der Referendare in der Stichprobe zurück. Auch das mangelnde Feedback zum fachdidaktischen Wirken der Referendare im Unterricht könnte ursächlich für die gefundene Diskrepanz sein. Dieser Befund macht umso deutlicher, dass für die Erfassung professioneller Kompetenzen von (angehenden) Lehrkräften zuverlässige Testinstrumente benötigt werden, die belastbare Schätzungen der zugrundeliegenden Kompetenz bieten können.

Bei der Erfassung professioneller Wissensinhalte (insbesondere PCK) hat sich vor allem ein Ansatz zur Konstruktion von Fremdbeurteilungsverfahren herausgebildet, der sog. Vignettentest (Brovelli, Bölsterli, Rehm & Wilhelm, 2013; König, 2015; Neuweg, 2015). Ein Vignettentest besteht aus komplexen Testaufgaben (Vignetten). In jeder Vignette wird eine authentische Unterrichtssituation beschrieben, die der Proband hinsichtlich einer konkreten fachdidaktischen Fragestellung in einem offenen (Kunter, Baumert et al., 2011) oder geschlossenen Antwortformat (Tepner & Dollny, 2014) beantworten soll. Dabei versetzt sich der Proband in die Rolle der fiktiven Lehrkraft, die in der Unterrichtssituation dargestellt wird. Hierbei liegt die Annahme zugrunde, dass durch die stellvertretende

Beurteilung des professionellen Verhaltens einer dritten Person (fiktive Lehrkraft in der Unterrichtsvignette) Rückschlüsse auf die professionelle Kompetenz bzw. die professionellen Wissensinhalte der Probanden gezogen werden können. Diese Annahme wird als sog. *advokatorischer Ansatz* bezeichnet (Heinzer & Oser, 2013; Oser et al., 2007).

Der *advokatorische Ansatz* betont die Notwendigkeit eines Bezugsrahmens für die Messung professioneller Kompetenzen von (angehenden) Lehrkräften. Professionelle Kompetenzen können sich demnach nur zeigen, wenn ein entsprechender professioneller Kontext vorgegeben wird anhand dessen sich die professionelle Reaktion bzw. das professionelle Handeln orientieren kann. Das Format der Unterrichtsvignette gibt einen authentischen Bezugsrahmen und Kontext vor (d. h. eine spezifische Unterrichtssituation), sodass sich die Probanden bei der Bearbeitung der Unterrichtsvignetten an diesem orientieren können. Dieses Testformat ist zur Erfassung von PCK vor allem deshalb sehr gut geeignet, da sowohl die Situiertheit als auch die Komplexität des Unterrichtsgeschehens abgebildet werden können (McNeill, González-Howard, Katsh-Singer & Loper, 2016).

Zu jeder Unterrichtsvignette wird eine fachdidaktisch relevante Fragestellung präsentiert, die sich zumeist darauf bezieht, wie der Proband an Stelle der fiktiven Lehrkraft den Unterricht fortsetzen würde (Riese & Reinhold, 2008). Der befragte Proband reagiert auf die Unterrichtssituation daher unter Rückgriff auf die eigenen praktischen Erfahrungen und Wissensbestände. Auf diese Art und Weise kann das implizite Wissen und Können der Probanden auf eine explizite Wissensebene angehoben werden, was eine empirische Abbildung der zugrunde liegenden Kompetenz ermöglicht (Hedtke, 2000). Durch das Verwenden von authentischen Unterrichtsvignetten wird daher der Tatsache Rechnung getragen, dass davon ausgegangen werden kann, dass es sich bei PCK um einen teilweise impliziten Wissensinhalt handelt.

In jüngster Zeit haben sich verschiedene Forschergruppen mit der Entwicklung von Testinstrumenten zur Erfassung professioneller Wissensinhalte von Lehrkräften beschäftigt: U. a. im Fachbereich Mathematik „Professionelle Kompetenz von Lehrkräften, kognitiv aktivierender Unterricht und die mathematische Kompetenz von Schülerinnen und Schülern“ (COACTIV, Kunter, Baumert et al., 2011), „Teacher Education and Development Study“ (TEDS; Buchholz et al., 2012), in

den Naturwissenschaften im Projekt „Naturwissenschaften unterrichten können“ (NUK, z. B. Brovelli, Bölsterli, Rehm & Wilhelm, 2014), „Professionswissen in den Naturwissenschaften“ (ProwiN; Tepner et al., 2012) und im Fachbereich Deutsch „Teacher Education and Development Study – Learning to Teach“ (TEDS-LT; Blömeke et al., 2013) und „Fachspezifische Lehrerkompetenzen“ (FALKO; Pissarek & Schilcher, 2015).

Bezogen auf den Kompetenzbereich „Professionelle Wahrnehmung“ haben Seidel, Blomberg und Stürmer (2010) ein valides Testinstrument vorgelegt, das videobasierte Vignetten als Testaufgaben verwendet (auch: Meschede, 2014). Südkamp, Möller und Pohlmann (2008) haben mithilfe der Computeranwendung „Simulierter Klassenraum“ die Genauigkeit von Leistungsbeurteilungen simulierter Schüler durch Lehramtsstudierende als einen Aspekt der diagnostischen Kompetenz von (angehenden) Lehrkräften untersucht. In der Computeranwendung „simulierter Klassenraum“ können die Probanden die Rolle einer fiktiven Lehrkraft einnehmen und mit den simulierten Schülern im Rahmen der Computeranwendung interagieren.

Abschließend kann festgestellt werden, dass das professionelle Wissen von (angehenden) Lehrkräften als bedeutungsvoller Kompetenzaspekt angesehen werden muss. So hat sich für die professionelle Wissensdomäne PCK für den Fachbereich Mathematik herausgestellt, „dass das fachdidaktische Wissen (PCK) größere Vorhersagekraft für den Lernfortschritt von Schüler(inne)n [als das Fachwissen (CK)] hat und maßgeblich die Unterrichtsqualität beeinflusst“ (Baumert & Kunter, 2011b, S. 182).

Die Entwicklung von PCK stellt daher ein zentrales Forschungsinteresse der Lehrerbildungsforschung dar, wobei der allgemeine Entwicklungsverlauf ebenso interessiert wie entsprechende Fördermöglichkeiten in Form von spezifischen Ausbildungsbausteinen in der Lehramtsausbildung (z. B. praktische Ausbildungsabschnitte). Um den Verlauf sowie Einflussfaktoren auf die Entwicklung von PCK zu untersuchen, werden standardisierte und erprobte Testinstrumente benötigt. Hierzu hat sich vor allem der Ansatz des vignettenbasierten Testens bewährt. Dieses Testformat trägt durch die Verwendung von situationsbezogenen komplexen Testaufgaben dazu bei, dass auch implizite Anteile von PCK erfasst werden können. Ergebnisse zum Einsatz von vignettenbasierten Testverfahren beschrän-

ken sich allerdings momentan (noch) auf die Domänen Mathematik und die Naturwissenschaften.

6 Lesekompetenz und Leseprozess

Eingangs wurde berichtet, dass viele Kinder und Jugendliche in Deutschland eine eingeschränkte Lesekompetenz aufweisen (Bos, Tarelli et al., 2012; Hohn et al., 2013). Ein Ziel dieser Studie ist daher das fachdidaktische Wissen und Können zu untersuchen, über das (angehende) Deutschlehrkräfte zur nachhaltigen Förderung der Lesekompetenz ihrer Schüler verfügen sollten. Dazu soll in diesem Kapitel zunächst der Begriff Lesekompetenz erläutert werden. Anschließend werden Ergebnisse der kognitionspsychologischen Forschung zu den mentalen Abläufen beim Leseprozess präsentiert. Zuletzt werden verschiedene Determinanten zum Auf- und Ausbau der Lesekompetenz vorgestellt, an die sich Ansatzpunkte zur Förderung anschließen.

6.1 Lesekompetenz: Begriff und Relevanz

Lesen zu können stellt in unserer heutigen Gesellschaft ein wichtiges Kulturwerkzeug dar und gilt als wesentliche Schlüsselkompetenz für die erfolgreiche Teilnahme am beruflichen und gesellschaftlichen Leben (Artelt, Stanat, Schneider & Schiefele, 2001).

Im schulischen Kontext zählt das Lesen zu den fachübergreifenden Kompetenzen, die die Schüler im Laufe ihrer Schulzeit erwerben (sollten): Lesen zu können betrifft nicht nur den Deutsch- bzw. Leseunterricht sondern ist ebenso bedeutsam für alle anderen Schulfächer, denn auch hier wird in der Regel mit Fachliteratur gearbeitet. Selbst im Mathematikunterricht wird das Lesen für das Verständnis von textbasierten Aufgaben benötigt. Daher kann konstatiert werden, dass die Fähigkeit des (informatischen) Lesens den Schulerfolg von Schülern in allen Unterrichtsfächern maßgeblich beeinflusst (Artelt et al., 2001). Ein zentrales Ziel der institutionalisierten Bildung in Deutschland muss daher sein, allen Schülern die Fähigkeit des Lesens in einem ausreichenden Maß durch den schulischen Unterricht zu vermitteln (Artelt & Dörfler, 2010).

Im Rahmen der ersten Erhebungswelle der internationalen Schulleistungsuntersuchung PISA wurde der Begriff *Lesekompetenz* geprägt und in den deutschen

Sprachraum etabliert (Artelt et. al., 2001). Unter Lesekompetenz ist folgendes zu verstehen:

„Der Begriff [Lesekompetenz] umfasst das Lesen als Ganzes, also gleichermaßen das, was in der Erwerbsphase angeeignet wird wie das automatisierte Lesen unterschiedlicher Lektüre (literarische wie nicht-literarische Texte; außerdem sog. „nicht-kontinuierliche“ Texte wie Tabellen und Diagramme). Dabei meint Kompetenz gleichermaßen Fähigkeiten und Fertigkeiten, schließt eine „(zeitlich) relativ stabile, individuell unterschiedliche Disposition“ ein (Grobe, 2002). Im Hinblick auf unterschiedliche Lektürearten werden das informatorische und das literarische Lesen unterschieden. Zum informatorischen Lesen gehört vor allem die Fähigkeit zur raschen Wahrnehmung, die Fähigkeit zeitökonomisch nach relevanten Informationen zu suchen, Teilaspekte in die Gesamtinformation wieder einzufügen und sie reflektieren und bewerten zu können (Hurrelmann, 2002). Dieser Begriff von Lesekompetenz steht bei den internationalen Vergleichsstudien PISA und IGLU im Vordergrund“ (Kliwer & Pohl, 2006).

Der Begriff Lesekompetenz ist also als ein Konglomerat an unterschiedlichen (Teil-)Fähigkeiten zu verstehen, die sich sowohl auf die technischen Fertigkeiten beim Lesen und die unterschiedlichen Teilprozesse beim Leseprozess beziehen, als auch auf die zuverlässige Informationsentnahme aus verschiedenen Textformaten.

6.2 Der Leseprozess

Lesen bezeichnet den Prozess der geistigen Verarbeitung eines Textes, der darauf abzielt dessen Bedeutung zu erfassen. Die unterschiedlichen mentalen Abläufe beim Leseprozess wurden von verschiedenen Autoren umfangreich beforscht und beschrieben (Kintsch & van Dijk, 1978; Richter & Christmann, 2002; Zwaan & Radvansky, 1998). Unter dem Begriff *Textverstehen* werden alle kognitiven Vorgänge, die an der Aufnahme, Organisation, Transformation, Speicherung, Reaktivierung und Reproduktion von Textinformationen beteiligt sind, zusammengefasst (Christmann, 2006).

Es besteht ein allgemeiner Konsens darüber, dass es sich beim Lesen nicht um eine passive Informations- und Bedeutungsaufnahme handelt, sondern dass Lesen als ein aktiver und konstruktiver Vorgang verstanden werden kann und muss (vgl. Christmann & Groeben, 1999). Bezüglich des Leseprozesses wird ein System unterschiedlicher mentaler Operationen auf verschiedenen Hierarchieebenen postuliert, das die Informationsentnahme aus und das Verständnis von Texten ermöglicht (Kintsch & van Dijk, 1978). Zu den hierarchieniedrigen Prozessen zählen die Wortidentifikation sowie die Bildung von lokalen Kohärenzen⁶, die bei geübten Lesern automatisch ablaufen. Zu den hierarchiehöheren Prozessen zählt die globale Kohärenzbildung⁷, die ein höheres Ausmaß an Konzentration auf Seiten des Lesers erfordert. Je stärker die Prozesse der hierarchieniedrigen Ebenen automatisiert ablaufen, über desto mehr kognitive Kapazität verfügt der Leser für das Textverstehen.

Es wird angenommen, dass die Informationsverarbeitung beim Leseprozess auf drei unterschiedlichen Repräsentationsebenen abläuft (Vgl. Christmann, 2006; vgl. Christmann, 2006; Schiefele, Artelt, Schneider & Stanat, 2004). Weiterhin werden beim und im Anschluss an das Lesen vom jeweiligen Vorwissen abhängige Inferenzen gebildet, die eine Integration von Textinformationen in die bisherigen Wissensbestände ermöglichen (Graesser, Singer & Trabasso, 1994).

⁶ Lokale Kohärenzbildung meint die (Sinn-)Verknüpfung zwischen einzelnen Wörtern, Wortgruppen oder aufeinander folgenden Sätzen.

⁷ Globale Kohärenzbildung meint die (Sinn-)Erfassung des gesamten Textes, das Lesen auf Textebene.

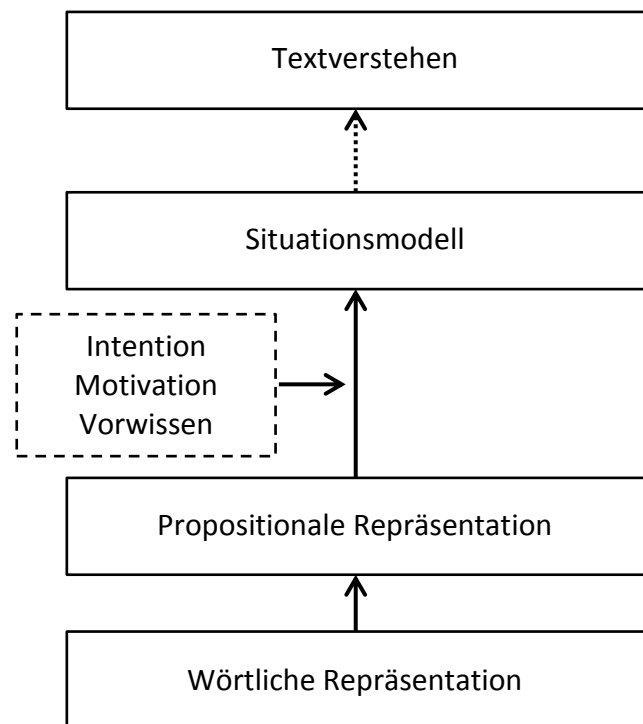


Abbildung 4. Vereinfachtes Modell des Textverstehens in Anlehnung an Kintsch (1992) und Artelt und Dörfler (2010).

Abbildung 4 zeigt ein vereinfachtes Modell des Textverstehens. Auf der untersten Verarbeitungsebene wird eine wörtliche Repräsentation des Textes gebildet. Dabei werden der genaue Wortlaut sowie graphemische, lexikalische und syntaktische Eigenschaften des Textes abgebildet. Die wörtliche Repräsentation des Textes wird nur kurzfristig im Arbeitsgedächtnis aufrechterhalten.

Auf der zweiten Verarbeitungsebene wird die semantische Bedeutung der Wörter erfasst, jedoch nicht (mehr) die exakte Wortfolge und Syntax. Zur Erfassung des semantischen Gehalts der Wörter und Sätze werden sog. Propositionen gebildet. Unter Propositionen werden komplexe Symbole verstanden, die Aussagen über den Gegenstand des Textes darstellen (Schnotz & Dutke, 2004).

Schließlich wird auf der obersten Repräsentationsebene ein Situationsmodell gebildet. Dabei findet eine integrative Verknüpfung des im Text geschilderten Sachverhalts bzw. der Situationsinformation und des Vorwissens des Lesers im Arbeitsgedächtnis statt. Bei der Bildung von Situationsmodellen spielen auch die Motivation und die Intention des Lesers eine bedeutsame Rolle, da diese den In-

formationsintegrationsprozess ebenfalls beeinflussen. Damit eine Integration der vorliegenden Textinformationen und dem Vorwissen gelingt, sind hierarchiehöhere Leseprozesse notwendig.

Wenn die Informationen aus dem Text – die auf diesen drei Ebenen repräsentiert werden – erfolgreich integriert und mit dem eigenen Vorwissen verknüpft werden können, kann davon ausgegangen werden, dass der Leser den Text verstehen kann.

6.3 Empirische Evidenzen zur Lesekompetenz deutscher Schüler

Mittlerweile liegen aus den verschiedenen nationalen und internationalen Schulleistungsuntersuchungen, die in Deutschland durchgeführt wurden, umfangreiche Ergebnisse zur Lesekompetenz von Grund- bzw. Sekundarstufenschülern vor.

Bezüglich der Primarstufe hat die IGLU-Studie gezeigt, dass 15.4 % der Grundschüler in Deutschland nicht die Lesekompetenzstufe III erreichen, was auf eine unzureichend ausgebildete Lesekompetenz hinweist (Bos, Tarelli et al., 2012). Insbesondere für diese Schülergruppe sind Lernschwierigkeiten in der Sekundarstufe in allen Unterrichtsfächern zu erwarten. Dieses Ergebnis ist allerdings vergleichbar zu anderen europäischen Staaten. Insgesamt weisen die Viertklässler mit $M = 541$ Punkten eine durchschnittliche Leseleistung auf, die sich im oberen Drittel des internationalen Vergleichsrahmens befindet (Bos, Tarelli et al., 2012). Die befragten Viertklässler in Deutschland berichten im Allgemeinen ein aktives Leseverhalten sowie eine positive Einstellung zum Lesen. Allerdings ist es als problematisch zu bezeichnen, dass sich die Leseleistung zwischen Schülern mit und ohne Migrationshintergrund deutlich unterscheidet und zwar dahingehend, dass Schüler mit Migrationshintergrund im Durchschnitt eine schlechtere Leseleistung aufweisen. Zudem kann festgestellt werden, dass Kinder aus sozial schwächeren Familien unterdurchschnittliche Leseleistungen aufweisen (vgl. Artelt & Dörfler, 2010).

Bezüglich der Sekundarstufe I hat die PISA-Studie gezeigt, dass sich die mittlere Leseleistung der 15-jährigen Schüler in Deutschland mit $M = 497$ Punkten ($SD = 95$) nicht statistisch bedeutsam vom OECD-Durchschnitt unterscheidet ($M = 500$,

$SD = 100$) (Naumann, Artelt, Schneider & Stanat, 2010). Allerdings weisen 18.5 % der befragten Schüler nur das Lesekompetenzniveau von Stufe Ia oder darunter auf. Ein äußerst geringer Anteil an Schülern von 7.6 % weist ein Kompetenzniveau von Stufe IV oder höher auf. Besorgniserregend ist also vor allem der große Prozentsatz an schwachen Lesern.

Daneben ist eine sehr große Leistungsstreuung bezüglich der Leseleistung zu beobachten; insbesondere Jugendliche mit Migrationshintergrund weisen dabei eine schwache Leseleistung auf. Weiterhin können deutliche Geschlechtsunterschiede dahingehend beobachtet werden, dass Mädchen eine durchschnittlich höhere Lesekompetenz aufweisen als Jungen. Der beobachtete Geschlechterunterschied geht auch mit einer unterschiedlich ausgeprägten Motivation einher; Mädchen berichten eine stärker ausgeprägte Lesemotivation als Jungen (vgl. Artelt & Dörfler, 2010). Ebenfalls kann konstatiert werden, dass sich seit Beginn des systematischen Bildungsmonitorings in Deutschland der Anteil an Schülern, die eine schwache Leseleistung aufweisen, nicht statistisch bedeutsam verändert hat (Klieme et al., 2010).

6.4 Determinanten von Lesekompetenz

Um die Lesekompetenz von Schülern im Unterricht nachhaltig zu fördern, sind fundierte Kenntnisse der Determinanten der Entwicklung von Lesekompetenz notwendig; an diesen Determinanten können entsprechende Förderprogramme und Trainings für Schüler orientiert werden (Rosebrock & Nix, 2014). Empirische Evidenzen zur Lesekompetenz in der Sekundarstufe I zeigen, dass vor allem die Vermittlung und Förderung (meta-)kognitiver Lesestrategien und die Steigerung der Lesemotivation effektive Maßnahmen zur Erhöhung der Lesekompetenz von Schülern darstellen (Artelt & Dörfler, 2010; Streblow, 2004; Caretti, Caldarola, Tencati & Cornoldi, 2014). Daneben haben Ansätze zur Steigerung und Festigung technischer Lesefertigkeiten stark an Bedeutung gewonnen (Scherf, 2013).

Im Folgenden werden Ansätze zur Förderung von Lesekompetenz, die im Rahmen des schulischen Unterrichts realisiert werden können, vorgestellt. Entsprechende fachdidaktische Vorgehensweisen zur Umsetzung dieser Fördermaßnahmen werden unter 7.1 erläutert.

6.4.1 Basale Lesefertigkeiten

Unter *basalen Lesefertigkeiten* werden verschiedene technische Lesefertigkeiten verstanden. Gut ausgebildete basale Lesefertigkeiten erleichtern die hierarchiehöheren Leseprozesse und damit einhergehend das Textverstehen (vgl. Rosebrock & Nix, 2014).

Lesegeschwindigkeit meint die Geschwindigkeit mit der Texte gelesen werden können. Die Lesegeschwindigkeit weist einen positiven Zusammenhang mit dem Textverstehen auf (Jenkins, Fuchs, van den Broek, Espin & Deno, 2003; Venendaal, Groen & Verhoeven, 2015). Liest eine Person zu langsam, können die gelesenen Informationen nicht ausreichend lange im Arbeitsgedächtnis präsent gehalten werden, um diese mit den vorherigen Textinformationen und dem eigenen Vorwissen zu verknüpfen. Gelingt dieser Integrationsprozess nicht bzw. nicht ausreichend, wird der Leser mit Schwierigkeiten beim Textverstehen konfrontiert.

Lesegenauigkeit bezieht sich auf die Fähigkeit des Lesers Wörter möglichst exakt zu dekodieren und auftretende Lesefehler korrigieren zu können. Eine mangelnde Genauigkeit beim Dekodieren der gelesenen Wörter kann mit einer Beeinträchtigung des Textverstehens einhergehen, vor allem wenn es durch fehlerhaftes Dekodieren zu einer Bedeutungsveränderung der gelesenen Wörter kommt (vgl. Artelt & Dörfler, 2010).

Automatisierung bezieht sich auf das Ausmaß in dem der Prozess des Dekodierens von Wörtern automatisiert abläuft. Nur wenn der Dekodierungsprozess beim Lesen in ausreichendem Maße automatisiert verläuft, stehen ausreichend kognitive Ressourcen zur weiteren Informationsverarbeitung zur Verfügung, die das Textverstehen überhaupt erst ermöglichen (LaBerge & Samuels, 1974).

Sinnbetontes Lesen meint die Fähigkeit, den Lesefluss an die Intonation der gesprochenen Sprache anzunähern. Durch sinnbetontes Lesen wird der Text schon beim Lesen in zusammenhängende Teilabschnitte gegliedert, was die Erfassung der Bedeutung des Textes erleichtern kann (Rosebrock & Nix, 2014).

Diese vier basalen Lesefertigkeiten – Lesegeschwindigkeit, Automatisierung, Genauigkeit, sinnbetontes Lesen – werden auch unter dem Begriff *Leseflüssigkeit* zusammengefasst. Die Förderung der Leseflüssigkeit (*fluency*) hat sich im anglo-amerikanischen Sprachraum bereits seit den 1970er Jahren als effiziente Maß-

nahme zur systematischen Förderung der Lesekompetenz von Schülern etabliert (Rosebrock & Nix, 2014). Im deutschen Sprachraum wurde hingegen lange Zeit angenommen, dass sich das Erlernen und die Festigung der basalen Lesefähigkeiten auf die Primarstufe beschränkt („learning to read“). Daher wurde entsprechend für den Beginn der Sekundarstufe ein ausreichendes Ausmaß an Lesekompetenz vorausgesetzt, was ein textbasiertes Lernen aus (Fach-)Texten ermöglicht („reading to learn“) (Rosebrock & Nix, 2014). Wie einschlägige empirische Befunde belegen, kann hiervon jedoch nicht zuverlässig ausgegangen werden.

Gelingt es Kindern und Jugendlichen nicht, ein ausreichendes Maß an Leseflüssigkeit zu erwerben, wirkt sich dies negativ auf ihre weitere Leseentwicklung aus (Garbe, Holle & Salisch, 2006). So haben Klicpera und Gasteiger-Klicpera (1993) längsschnittlich gezeigt, dass Kinder, die in der zweiten und dritten Klasse eine eingeschränkte Leseflüssigkeit aufweisen, diesen Rückstand nicht ohne zusätzliche Fördermaßnahmen aufholen können. Insbesondere im Bereich der Hauptschule wird vielen Schülern zu Beginn der Sekundarstufe eine unzureichende Leseflüssigkeit bescheinigt: Über 50 % der befragten Sechstklässler einer Hauptschule wiesen in einer repräsentativen Erhebung unterdurchschnittliche Werte bezüglich basaler Lesefertigkeiten auf (Gold, Nix, Rieckmann & Rosebrock, 2010).

6.4.2 Lesemotivation

Lesemotivation bezeichnet das Ausmaß des Wunsches oder der Absicht einer Person, in einer spezifischen Situation einen bestimmten Text zu lesen (Möller & Schiefele, 2004). Die von Deci und Ryan (1985) getroffene Unterscheidung in extrinsische und intrinsische Aspekte von Motivation trifft dabei auch für den Bereich der Lesemotivation zu: Unter einer intrinsischen Lesemotivation wird eine Leseabsicht verstanden, die durch ein persönliches Interesse an einem Thema oder durch die Freude an der Tätigkeit des Lesens begründet ist. Unter einer extrinsischen Lesemotivation versteht man eine Leseabsicht, die im Fall von Schülern, z. B. darauf abzielt, eine gute Note oder ein Lob des Lehrers zu erhalten. Extrinsische Lesemotivation wird demnach von außen motiviert.

Eine hohe intrinsische Lesemotivation geht mit einer höheren Lesekompetenz einher: So berichtet die PISA-Studie einen korrelativen Zusammenhang von $r = 0.46$ ($p \leq 0.05$) zwischen der intrinsischen Motivation (erfasst über den Indi-

kator „Lesefreude“) und der Lesekompetenz (Artelt, Naumann & Schneider, 2010, S. 104). Dieser Zusammenhang ist dadurch zu erklären, dass eine höhere intrinsische Lesemotivation zu einer höheren (außerschulischen) Lesemenge führt, was sich (durch das wiederholte „Üben“ der Tätigkeit Lesen) wiederum förderlich auf das Textverstehen auswirkt (vgl. Pfost, Dörfler & Artelt, 2010; Schön, 2002).

6.4.3 Lesestrategien

Lesestrategien bezeichnen in verschiedenen (Lese-)Situationen einsetzbare und zielgerichtete Vorgehensweisen zur Erschließung von Texten, die automatisiert ablaufen können, jedoch auch bewusstseinsfähig sind (Spinner, 2004). Ein adäquater Einsatz von Lesestrategien ermöglicht es, den eigenen Lese- und Verstehensprozess besser zu steuern und zu überwachen. Lesestrategien können in kognitive und metakognitive Strategien unterschieden werden (siehe auch Streblow, 2004):

Kognitive Lesestrategien beziehen sich auf die Organisation (z. B. Übersichten ausarbeiten, Zusammenfassungen schreiben), Wiederholung (z. B. mehrmaliges Lesen von Textpassagen) und Elaboration (z. B. praktische Beispiele suchen) des Gelesenen.

Metakognitive Lesestrategien beziehen sich auf das Planen (z. B. das Setzen von Zielen), Überwachen (des eigenen Lesevorgangs) und die Kontrolle des Leseprozesses (z. B. Verständnisschwierigkeiten erkennen und entsprechend eingreifen).

Die PISA-Studie berichtet für den Lesestrategie-Indikator „Verwendung von Kontrollstrategien“ einen korrelativen Zusammenhang zur Lesekompetenz von $r = 0.24$ ($p \leq 0.05$) und für den Lesestrategie-Indikator „Lernstrategiewissen“ einen korrelativen Zusammenhang zur Lesekompetenz von $r = 0.55$ ($p \leq 0.05$) (Artelt et al., 2010, S. 104).

7 Professionelles Wissen von (angehenden) Deutschlehrkräften

Für den Bereich Mathematik deuten bereits zahlreiche Forschungsbefunde darauf hin, dass professionelle Wissensinhalte von (angehenden) Lehrkräften bedeutsam für die Unterrichtsqualität sind (Blömeke, Kaiser & Lehmann, 2008; Kunter, Baumert et al., 2011). Herzmann und König (2016) stellen allerdings fest: „Inwieweit ... Fachwissen und fachdidaktisches Wissen von ... Lehrkräften sprachlicher Fächer bedeutungsvoll für ihren Unterricht sind, ist eine weitgehend offene Forschungsfrage.“ (S.119). Im Fachbereich Deutsch ist daher „für den Bereich der Lesekompetenz [...] bisher empirisch kaum überprüft, inwieweit die Unterrichtsqualität und die Schülerleistungsentwicklung mit den Kompetenzen der Lehrkräfte in Zusammenhang steht.“ (McElvany & Schneider, 2009, S. 155).

Eine Ursache hierfür ist u. a., dass im wissenschaftlichen Diskurs zur Lese- und Literaturdidaktik die empirische Bildungsforschung keineswegs als selbstverständlich und notwendig angesehen wird, so wie es innerhalb anderer Fachdidaktiken, beispielsweise in den Naturwissenschaften, und von bildungspolitischer Seite aus der Fall ist (Hurrelmann & Groeben, 2006). Dies ist z. T. darauf zurückzuführen, dass die Tradition der Lese- und Literaturdidaktik in Deutschland auf einer hermeneutischen Herangehensweise an fachbezogene Problemstellungen beruht, die teilweise unvereinbar mit einer empirischen Herangehensweise erscheint. Um den gesellschaftlichen wie politischen Anforderungen gerecht(er) zu werden, fordern Vertreter der Deutschdidaktik zunehmend, „die Ansätze der empirischen Forschung im eigenen Feld fachspezifisch, unterrichtswissenschaftlich und methodologisch systematisch aufzubauen“ (Hurrelmann & Groeben, 2006, S. 12).

Die Fachdidaktik im Unterrichtsfach Deutsch gilt als „gering strukturierte Domäne“ (Blömeke et al., 2011), für die eine Ableitung konkreter professioneller Kompetenzen, die Lehrkräfte innerhalb ihres Studiums erwerben sollten, bisher ausblieb. In den letzten fünfzehn Jahren sind dessen ungeachtet einige Forschungsarbeiten zur Professionalisierung von Deutschlehrkräften entstanden, die eine qualitative Methodik (Bräuer, 2010; Scherf, 2013; Wieser, 2008) bzw. eine quantitative

Methodik (Bremerich-Vos & Dämmer, 2013; Pissarek & Schilcher, 2015) zugrunde legen.

7.1 Fachdidaktisches Wissen im Leseunterricht

Der Leseprozess als solcher sowie die beteiligten mentalen (Teil-)Fertigkeiten auf den verschiedenen Hierarchieebenen wurden bereits ausführlich beforscht und beschrieben. Darüber hinaus sind auch deutsche Schüler in der Primar- und Sekundarstufe ausführlich bezüglich ihrer Lesekompetenz untersucht worden. Daraus resultierten klare Empfehlungen darüber, welche Teilfertigkeiten die Schüler erwerben müssen, um erfolgreich lesen zu können, beispielsweise eine ausreichende Leseflüssigkeit. Ebenso wurden bereits zahlreiche Förderprogramme zur systematischen Förderung der Lesekompetenz im Rahmen des schulischen Unterrichts entwickelt und evaluiert (z. B. Rosebrock, Nix, Rieckmann & Gold, 2011).

Wie eingangs beschrieben ist die Befundlage bezüglich der professionellen Kompetenzen von Deutschlehrkräften als weitaus eingeschränkter zu bezeichnen. Es wurde bislang kaum empirisch untersucht, über welche professionellen Wissensinhalte Deutschlehrkräfte verfügen müssen, um die Lesekompetenz ihrer Schüler durch entsprechende Maßnahmen im Leseunterricht effizient zu fördern. Für die Domäne Mathematik liegen bereits gesicherte empirische Evidenzen darüber vor, dass insbesondere der professionelle Wissensinhalt PCK (mit verschiedenen inhaltlichen Wissensfacetten) positiv mit der Schülerleistungsentwicklung in Zusammenhang steht. Analog dazu kann ein entsprechender Wirkzusammenhang auch für Deutschlehrkräfte angenommen werden.

Die Untersuchung von PCK bei (angehenden) Deutschlehrkräften bezüglich des Leseunterrichts stellt daher einen zielführenden Ansatzpunkt zur Förderung der Lesekompetenz von Schülern dar: Angelehnt an das Determinationsmodell der Schulleistung von Harnischfeger und Wiley (1977) wird erwartet, dass es sich lohnt, professionelle Kompetenzen von Lehrkräften umfassend zu fördern (z. B. bezüglich des professionellen Wissens). Veränderungen der Kompetenzen auf Lehrerseite (hier: Zuwachs an Wissen und Fähigkeiten bzgl. der Förderung von Lesekompetenz) haben nach Harnischfeger und Wiley eine direkte Auswirkung auf die konkreten Lehrkraftaktivitäten im Unterricht und in Folge dessen auch auf die Schülerleistung (hier: Steigerung der Lesekompetenz der Schüler). Um aller-

dings evidenzbasiert Fördermaßnahmen für (angehende) Deutschlehrkräfte entwickeln zu können, sind fundierte Kenntnisse über die Struktur und die Entwicklung der zugrunde liegenden professionellen Wissensbestände – wie beispielsweise PCK – notwendig.

Bezüglich der in 6.4.1 bis 6.4.3 dargestellten Determinanten der Lesekompetenz ergeben sich für den schulischen Leseunterricht verschiedene fachdidaktische Ansätze, die zu einem Aufbau der Lesekompetenz der Schüler beitragen (können). Daher sollten (angehende) Deutschlehrkräfte idealerweise über ein entsprechendes Wissen und Können dieser fachdidaktischen Prinzipien und Vorgehensweisen verfügen.

7.1.1 Fachdidaktische Ansätze zur Steigerung basaler Lesefertigkeiten

Der Auf- und Ausbau basaler Lesefertigkeiten stellt wie oben beschrieben nicht nur eine fachdidaktische Aufgabe für Deutschlehrkräfte im Primarstufenbereich, sondern auch im Sekundarstufenbereich dar. Angelehnt an ein umfassendes Instrumentarium an evidenzbasierten Fördermöglichkeiten, das bereits seit den 1970er Jahren vorrangig in den USA entwickelt, evaluiert und eingesetzt wurde und wird, wurden auch im deutschen Sprachraum zunehmend Förderkonzepte zur Steigerung basaler Lesefertigkeiten in die deutschdidaktische Diskussion eingeführt. Zur Förderung der basalen Lesefertigkeiten im schulischen Unterricht gelten *Lautleseverfahren* sowie *Vielleseverfahren* als relevant.

Lautleseverfahren bezeichnen Methoden, bei denen die Schüler im Unterricht bekannte oder unbekannte Texte (halb-)laut vorlesen (Rosebrock & Nix, 2014). Von dem klassischen „Reihumlesen“ wird aus aktueller deutschdidaktischer Sicht allerdings abgeraten: Beim „Reihumlesen“ wird ein unbekannter Text in der Klasse ausgegeben, der von einem Schüler spontan laut vorgelesen werden soll (ohne vorheriges Üben und ohne dass der vorlesende Schüler den Inhalt des Textes kennt). Für den vorlesenden Schüler ist bei dieser Methode aufgrund der Spontaneität der Anforderung nicht mit einem Übungseffekt zu rechnen; zudem kann ein stockendes Vorlesen zu Verständnisschwierigkeiten bei den übrigen Schülern der Klasse führen. Weiterhin ergeben sich durch die Methode „Reihumlesen“ u. U. ungünstige (Lese-)Situationen und zwar dahingehend, dass schwache Leser leicht

stigmatisiert werden und starke Leser sich während der Aufgabe langweilen. Daneben belegen die Ergebnisse einschlägiger empirischer Studien, dass die Effekte dieser Methode als nicht eindeutig gewertet werden müssen (Opitz & Rasinski, 1998).

Lautleseverfahren können in das *wiederholte Lautlesen* und das *begleitete Lautlesen* unterschieden werden (siehe auch: Nix, 2011): Beim wiederholten Lautlesen handelt es sich um Methoden, bei denen die Schüler Texte so lange laut vorlesen, bis sie eine bestimmte Lesegeschwindigkeit erreicht haben. Dazu dient ein vordefinierter Wert an Wörtern pro Minute, den die Schüler erreichen sollen. Dabei überwacht ein Tutor oder der Lehrer den Übungsprozess. Angepasst an die vorherige Leseleistung der Schüler werden für die Leseübung leichte, mittelschwere oder schwierige Texte ausgewählt, um die Schüler auf ihrem individuellen Leseniveau adäquat zu fördern.

Beim begleiteten Lautlesen wird vor allem mit Lesemodellen gearbeitet. Hierbei werden beispielsweise Zweiergruppen aus einem schwachen und einem starken Leser gebildet, die sich gegenseitig Texte vor- bzw. gemeinsam („im Chor“) lesen; der starke Leser dient dann hinsichtlich der Lesegeschwindigkeit und der Betonung der Wörter als Lesevorbild. Im deutschen Sprachraum wurde bei Sechstklässlern einer Hauptschule erfolgreich die Methode des „Lautlese-Tandems“ eingesetzt, was zu einer signifikanten Verbesserung der Leseflüssigkeit der teilnehmenden Schüler führte (Nix, 2011).

Vielleseverfahren bezeichnen Methoden, die die Lesedauer der Schüler steigern sollen. Im schulischen Kontext werden Vielleseverfahren beispielsweise dahingehend umgesetzt, dass im Deutschunterricht feste, freie Lesezeiten vereinbart werden. In dieser Lesezeit lesen die Schüler selbst ausgewählte Bücher und Texte still für sich, wobei der Lehrer als Rollenmodell dient und selbst ein Buch bzw. einen Text liest. Die gelesenen Inhalte werden im Unterricht nicht thematisiert. Bei Vielleseverfahren wird angenommen, dass durch das vermehrte Ausüben der Tätigkeit „Lesen“ sowie durch die damit verbundene Erhöhung der Lesemenge ein positiver Effekt auf die Lesekompetenz zu erwarten ist (Rosebrock & Nix, 2014).

7.1.2 Fachdidaktische Ansätze zur Steigerung der Lesemotivation

Methoden zur Steigerung der Lesemotivation (auch: „Leseanimation“) zielen darauf ab, die Leselust der Schüler sowie deren Interesse an Büchern zu steigern. Im Idealfall tragen leseanimierende Methoden zum Aufbau intrinsischer Lesemotivation bei. Leseanimierende Ansätze im Deutschunterricht sind u. a. das Einrichten von Bücherkisten/Klassenbibliotheken, der Einsatz von Hörbüchern oder Buchvorstellungen von aktuellen Büchern/Lieblingsbüchern (Rosebrock & Nix, 2014).

Leseanimierende Methoden sind vor allem für Schüler geeignet, die über ein ausreichendes Maß an basalen Lesefertigkeiten verfügen und daher ohne größere Anstrengung sinnentnehmend lesen können. Für Schüler mit unzureichend ausgeprägten basalen Lesefertigkeiten können solche Verfahren zur Steigerung der Lesemotivation jedoch eine Überforderung darstellen und so eher einen gegenteiligen Effekt bewirken (z. B. Abschreckung oder Angst). Darüber hinaus sollte bei der Leseförderung mithilfe leseanimierender Verfahren darauf geachtet werden, dass sich Jungen und Mädchen hinsichtlich ihrer Leseinteressen unterscheiden (Philipp, 2013).

Garbe (2007) hat hierzu die bedeutsamsten langfristigen Geschlechterunterschiede beim Lesen zusammengestellt, die aus nationalen und internationalen Studien bekannt sind: Jungen und Mädchen unterscheiden sich zunächst hinsichtlich der Lesemenge und zwar dahingehend, dass Mädchen häufiger lesen als Jungen. Diese Beobachtung gilt ebenfalls für das Erwachsenenalter. Weiterhin präferieren Jungen und Mädchen unterschiedliche Lesestoffe: Jungen interessieren sich eher für Themen wie Kampf und Abenteuer, Herausforderung und Bewährung, Reise- und Heldengeschichten, wohingegen Mädchen eher von Beziehungs-, Liebes- und Tiergeschichten angesprochen werden. Zudem berichten Mädchen gegenüber Jungen ein höheres Ausmaß an Lesefreude und bezeichnen Lesen auch signifikant häufiger als eine favorisierte Freizeitbeschäftigung.

7.1.3 Fachdidaktische Ansätze zu Förderung des Einsatzes von Lesestrategien

Obwohl Lesestrategien (bei geübten Lesern) intuitiv angewendet werden (Rosebrock & Nix, 2014), hat es sich als sehr lohnenswert herausgestellt, Lesestrategien explizit im Unterricht zu vermitteln (Hattie, 2009). Das Einüben von Lesestrategie-

gien eignet sich vor allem für Schüler, die über ein ausreichendes Maß an Leseflüssigkeit verfügen, jedoch Probleme damit haben, die gelesenen Informationen sinnvoll in Beziehung zueinander zu setzen, relevante Informationen im Text zu identifizieren und/oder den Leitgedanken bzw. die Intention des Textes zu erfassen. Diese Schüler können einen Text zwar „durchlesen“, aber nicht sinnvoll aus diesem lernen (Kintsch, 1996). Schüler, die hierarchieniedrige Leseprozesse bereits gut beherrschen, die hierarchiehöheren Leseprozesse jedoch (noch) nicht, profitieren daher in besonderem Maße vom gezielten Einüben von Lesestrategien.

Für den deutschen Sprachraum liegen mittlerweile zahlreiche Lesestrategie-Programme vor, die sich für den Einsatz im Schulunterricht eignen (Überblick bei: Philipp & Schilcher, 2012). Zur nachhaltigen Vermittlung von Lesestrategien im Unterricht sind dabei drei didaktische Prinzipien unerlässlich (vgl. Rosebrock & Nix, 2014): Zunächst müssen die Lesestrategien in den deklarativen Wissensbestand der Schüler integriert werden. Dies kann darüber erfolgen, dass die Schüler die unterschiedlichen Lesestrategien und -techniken kennen lernen (z. B. „Was kann ich machen, wenn ich ein Wort nicht verstehe?“). Auf diese Art und Weise können im Deutschunterricht „Strategie-Kataloge“ angelegt werden. Es bietet sich dabei an, den Katalog am zeitlichen Ablauf des Leseprozesses zu orientieren, z. B. „Kapitel 1: Strategien vor dem Lesen“, „Kapitel 2: Strategien während dem Lesen“ und „Kapitel 3: Strategien nach dem Lesen“. Diese „zeitliche Anordnung“ der verschiedenen Lesestrategien trägt zu einer verbesserten Überschaubarkeit und Handhabbarkeit der Lesestrategien für die Schüler bei.

Wenn die Schüler die verschiedenen Lesestrategien kennen, ist es enorm wichtig, dass die Schüler üben, diese anzuwenden. Nur so kann die Verwendung von Lesestrategien in das „Verhaltensrepertoire“ der Schüler eingebettet werden. Zuletzt sollten die Schüler in der Lage sein, das umfangreiche Repertoire an Lesestrategien selbstständig und reflektiert beim Lesen einzusetzen. Hierbei werden vor allem die metakognitiven Lesestrategien gefordert.

Besondere Aufmerksamkeit verdient die Tatsache, dass es sich bei der nachhaltigen Vermittlung von Lesestrategien im Rahmen des schulischen Unterrichts nicht um eine isolierte Unterrichtseinheit handelt. Vielmehr handelt es sich bei der nachhaltigen Förderung des Einsatzes von Lesestrategien um einen langfristigen Prozess, der stetige Wiederholungen der Lesestrategien und Übungsphasen für die

Schüler beinhaltet, wenn eine dauerhafte Steigerung der Textverstehensleistungen der Schüler herbeigeführt werden soll (Artelt & Dörfler, 2010).

7.1.4 Diagnose von Lese- und Verstehensprozessen

Korrektes und zuverlässiges Diagnostizieren von Schülerleistungen gilt gemeinhin als zentrales Merkmal von Lehrerexpertise (Helmke, 2015) und als Voraussetzung zur angemessenen sowie lernförderlichen Gestaltung des Unterrichts (u. a. bezogen auf die Aufgabenauswahl, die Bewertung, usw.) (Helmke, Hosenfeld & Schrader, 2004). Dabei ist die diagnostische Kompetenz von Lehrkräften als Teil des professionellen Wissensinhalts PCK zu verstehen (Artelt, 2009). In empirischen Studien konnte daneben gezeigt werden, dass positive Zusammenhänge zwischen der diagnostischen Kompetenz der Lehrperson und den Schülerleistungen beobachtet werden können (Helmke, 2015).

Bezogen auf den Leseunterricht betrifft die Diagnosefähigkeit von Lehrkräften u. a. die korrekte Einschätzung des Niveaus, auf dem Schüler Texte lesen und verstehen können. Für eine effektive Förderung der Lesekompetenz sollten sowohl gelingende bzw. misslingende Lese- und Textverstehensprozesse korrekt diagnostiziert werden können als auch ausreichende Kenntnisse über notwendige Verstehensanforderungen von Texten vorliegen (Artelt, 2009). Ebenso ist ein ausreichendes Maß an Wissen über schwierigkeitsgenerierende Textmerkmale für eine effektive Förderung der Lesekompetenz bedeutsam (Bertschi-Kaufmann, 2007).

Zur Diagnose und Förderung der Lesekompetenz im schulischen Unterricht stehen Deutschlehrkräften verschiedene (Lese-)Aufgabentypen zur Verfügung, die immer einen „Doppelcharakter“ aufweisen (vgl. Lindauer & Schneider, 2007): Einerseits geht es bei Leseaufgaben darum, den gegebenen Primärtext zu verstehen und andererseits um die Bearbeitung bzw. Beantwortung der zugehörigen Verständnisfrage(n). Es gibt einerseits Leseaufgaben, die vorrangig auf eine Steigerung der Lesekompetenz der Schüler abzielen („Lernaufgabe“); hier werden konstruierte Anforderungssituationen präsentiert, die die Lesekompetenz der Schüler ausbilden, aufbauen und festigen sollen. Andererseits gibt es Leseaufgaben, die darauf abzielen das erreichte Leseniveau der Schüler zu überprüfen („Prüfungsaufgaben/Testaufgaben“).

Leseaufgaben weisen unterschiedliche Schwierigkeiten auf. Diese wird von drei Faktoren bedingt: (1) Dem Primärtext, (2) der Aufgabenstellung und (3) der Komplexität der ausgelösten kognitiven Prozesse (Lindauer & Schneider, 2007).

Schwierigkeitsgenerierende Merkmale des Primärtextes sind u. a. der formale Schwierigkeitsgrad (z. B. Schriftgröße, Satzspiegel), die Sprachkultur (z. B. Wortlänge, Wortwahl) und die Handlung/Erzählweise (z. B. Art/Anzahl der Spielorte) (Riss, 2005; Willenberg, 2005).

Bezüglich der Aufgabenstellung ist zwischen einem offenen und geschlossenen Antwortformat zu unterscheiden: Beim offenen Antwortformat müssen die Schüler die Lösung selbst generieren und dabei Format und Formulierung selbst wählen. Beim geschlossenen Antwortformat bieten sich für Leseaufgaben Multiple-Choice Aufgaben, Umordnungsaufgaben oder Durchstreichaufgaben an. Bei der Wahl des Antwortformates für eine Leseaufgabe ist zu bedenken, dass beim offenen Antwortformat durch das freie Formulieren einer Antwort nicht nur das Lesen, sondern auch die Schreibkompetenzen der Schüler geprüft werden: So können zwei Schüler einen Text zwar „gleich gut“ verstanden haben, ihre Antwort jedoch nicht „gleich gut“ schriftlich ausformulieren (vgl. Lindauer & Schneider, 2007).

Die kognitiven Anforderungen, die aus verschiedenen Leseaufgaben resultieren, können wie folgt kategorisiert werden (Lindauer & Schneider, 2007; angelehnt an: Narciss & Proske, 2001): Isolieren (Informationen, die zur Beantwortung der Frage relevant sind, müssen identifiziert werden), Erinnern (Neugelesenes muss mit bestehendem Wissen verknüpft werden), Transformieren (Gelesenes muss in neuer/veränderter Form wiedergegeben werden), Klassifizieren (Gemeinsamkeiten und Unterschiede zwischen den Figuren/Handlungen im Text finden), Argumentieren (Textinformationen in Argumentationskette anordnen) und Interpretieren (zuvor Geschehenes deuten/reflektieren und anhand dessen Vorhersagen machen). Die kognitiven Anforderungen und der Komplexitätsgrad steigen dabei zunehmend an.

Bezüglich der diagnostischen Kompetenz von Lehrkräften bezogen auf die Lesekompetenz eigener Schüler liegen aus den Ergebnissen des systematischen Bildungsmonitorings in Deutschland folgende Befunde vor: Es kann davon ausge-

gangen werden, dass Lehrkräfte die Leseleistungen ihrer Schüler häufig überschätzen (Niveauelemente), wohingegen das Bilden einer „Leseleistungs“-Rangfolge der Schüler innerhalb einer Klasse besser gelingt (Rangkomponente). Darüber hinaus liegen bezüglich der diagnostischen Kompetenz von Deutschlehrkräften große interindividuelle Unterschiede vor; d. h. dass sich die Güte der diagnostischen Urteile von Deutschlehrkräften in beträchtlichem Ausmaß unterscheidet (vgl. Artelt, 2009).

7.1.5 Fazit zu fachdidaktischen Ansatzpunkten zur Förderung von Lesekompetenz

(Angehenden) Deutschlehrkräften stehen verschiedene fachdidaktische Ansatzpunkte zur Verfügung, um im schulischen Unterricht zu einer Steigerung der Lesekompetenz der eigenen Schüler beizutragen. Ein ausreichendes Maß an diagnostischer Kompetenz ist dabei als Voraussetzung für den effizienten Einsatz von fachdidaktischen Ansätzen zur Steigerung der Lesekompetenz anzusehen. Die Lehrkraft muss zunächst korrekt diagnostizieren können, wie gut die eigenen Schüler lesen können und, falls Einschränkungen bezüglich der Lesekompetenz beobachtet werden, worauf diese zurückzuführen sind. Hierzu sollten die Lehrkräfte über fundiertes Wissen über Typen von Leseaufgaben und schwierigkeitsgenerierende Textmerkmale verfügen, um Leseaufgaben auswählen zu können, die zur Diagnose von Lese- und Verstehensleistungen geeignet sind. Daneben benötigen Deutschlehrkräfte fundierte Kenntnisse über die verschiedenen Ebenen des Leseprozesses, um eingeschränkte Teilfertigkeiten beim Lesen identifizieren zu können (z. B. Einschränkung bei hierarchiehöheren Leseprozessen).

Zur Förderung der basalen Lesefertigkeiten eignen sich Lautleseverfahren, die in den Unterricht eingebunden werden können. Vielleseverfahren können ein generelles Interesse am Lesen fördern. Zur Steigerung der Lesemenge im Allgemeinen können im schulischen Unterricht auch leseanimierende Verfahren zur Steigerung der (intrinsischen) Motivation der Schüler eingesetzt werden. Ein wünschenswerter Nebeneffekt von leseanimierenden Verfahren wäre, dass durch eine Steigerung der intrinsischen Lesemotivation auch der Umfang an außerschulischen Leseaktivitäten gesteigert wird. Beim Einsatz von Leseanimationsverfahren gilt es jedoch zu beachten, dass sich diese Verfahren nur für Schüler eignen, die bereits in aus-

reichendem Maße sinnentnehmend lesen können und dass sich die Leseinteressen von Jungen und Mädchen stark unterscheiden. Zur Förderung von hierarchiehöheren Leseprozessen (z. B. Sinnentnahme aus Texten) eignen sich insbesondere Verfahren, die auf eine Etablierung und Festigung des Einsatzes von Lesestrategien abzielen.

Zusammenfassend kann daher davon ausgegangen werden, dass die hier beschriebenen fachdidaktischen Ansatzpunkte als vielversprechend für die systematische Steigerung der Lesekompetenz im schulischen Unterricht angesehen werden können. Daher wäre es wünschenswert, wenn (angehende) Deutschlehrkräfte über das fachdidaktische Wissen und Können verfügen bzw. entsprechend ausgebildet und geschult werden würden, um diese fachdidaktischen Prinzipien adäquat im schulischen Unterricht einzusetzen.

7.2 Empirische Evidenzen zum fachdidaktischen Wissen (angehender) Lehrkräfte im Leseunterricht

Aus Kapitel 7.1 wird ersichtlich, dass (angehende) Deutschlehrkräfte bezüglich des Leseunterrichts über verschiedene fachdidaktische Wissensinhalte verfügen sollten, um die Lesekompetenz der Schüler optimal zu fördern. Hinsichtlich wichtiger (Teil-)Kompetenzen (insbesondere PCK) von Deutschlehrkräften für den Leseunterricht liegt allerdings bisher eine eingeschränkte empirische Befundlage vor. Im Folgenden wird PCK bezogen auf den Leseunterricht als *lesedidaktisches Wissen* bezeichnet. *Lesedidaktisches Wissen* wird – basierend auf obigen Darstellungen zu PCK – als professioneller Wissensinhalt von Deutschlehrkräften mit impliziten und expliziten Anteilen konzeptualisiert, der aus einer Verschmelzung von pädagogisch-psychologischem Wissen und dem Fachwissen bezüglich des Leseprozesses und der Lesedidaktik resultiert. In dem Modell „Konzepte des ‚Lehrerwissens‘“ ist *lesedidaktisches Wissen* dem Wissensbereich 2 zugeordnet (siehe Abbildung 2).

Im Bereich der empirisch-quantitativen Forschung sind vor allem zwei Studien zu nennen, die sich mit professionellen Wissensinhalten – und im engeren Sinne mit PCK – von Deutschlehrkräften beschäftigt haben:

Die TEDS-LT-Studie (Teacher Education and Development Study: Learning to Teach) (Blömeke et al., 2011; Blömeke et al., 2013; Bremerich-Vos & Dämmer, 2013) untersucht Lehramtsstudierende mit dem Fach Deutsch bezüglich deren Wissenszuwachs mit Aufgaben zu CK und PCK zu zwei Messzeitpunkten. Es wurden sowohl geschlossene als auch offene Fragen vorgelegt. Bezüglich PCK wurden zwei Dimensionen – Literaturdidaktik und Sprachdidaktik – operationalisiert⁸. Die Skala zur Literaturdidaktik weist eine EAP-Reliabilität von 0.62 (Bremerich-Vos & Dämmer, 2013, S. 59). Ferner zeigt sich ein marginal signifikanter Anstieg bezüglich der Skala zur Literaturdidaktik vom ersten Messzeitpunkt ($N = 337$; 3.-5. Semester) zum zweiten Messzeitpunkt ($N = 288$; 6.-8. Semester), $p = 0.051$ (Bremerich-Vos & Dämmer, 2013, S. 70).

Die FALKO-Studie („Fachspezifische Lehrerkompetenzen“) legte ein Testinstrument zur Erfassung von CK und PCK von Lehrkräften im Deutschunterricht vor. Die Testaufgaben beziehen sich auf verschiedene Bereiche des Deutschunterrichts wie Orthographie, Textproduktion, Literatur und Lesen⁹ (Pissarek & Schilcher, 2015). Bisher wurden erste Ergebnisse, u. a. zur internen Konsistenz der Leseskala, veröffentlicht. Es handelt sich bezüglich des Aufgabenbereichs zu PCK hauptsächlich um Testaufgaben mit offenem Antwortformat. Diese wurden durch zwei unabhängige Rater codiert. Ein Wert von $\kappa = 0.7$ ist als ein Indikator für eine gute Übereinstimmung zwischen den Ratern anzusehen (Pissarek & Schilcher, 2015, S. 336). Die Testaufgaben wurden an einer Stichprobe von Lehrkräften (Gymnasium und Realschule) sowie Referendaren im letzten Ausbildungsabschnitt und Studierenden aus Examenskursen validiert ($N = 103$). Erste Ergebnisse zeigen eine akzeptable interne Konsistenz der Aufgabengruppe zu PCK von $\alpha = 0.67$ (Pissarek & Schilcher, 2015, S. 336). Außerdem konnten die Autoren einen Zusammenhang zwischen PCK und CK zeigen ($r = 0.52$, $p < 0.01$) (Pissarek & Schilcher, 2015, S. 337). Ein weiteres Ergebnis dieser Studie ist, dass die Inhalte der Aufgaben zu PCK von den Probanden als signifikant relevanter für die Ausbildung von Lehrkräften eingeschätzt wurden als die Aufgaben zum CK.

⁸ Es werden hier nur die Ergebnisse der Dimension „Literaturdidaktik“ präsentiert, da in der vorliegenden Studie vorrangig das fachdidaktische Wissen im Leseunterricht interessiert.

⁹ Es werden hier nur die Ergebnisse zum Aufgabenbereich „Fachdidaktik“ präsentiert, da in der vorliegenden Arbeit vorrangig das fachdidaktische Wissen von Deutschlehrkräften interessiert.

Bezüglich des Wissensbereichs CK haben Rausch, Matthäi und Artelt (2015) untersucht, ob (Fach-)Wissen über das Textverstehen einen Einfluss auf die Urteilsgüte und die Einschätzungen von Schülerleistungen hat. Hier konnten allerdings keine substanziellen Zusammenhänge zwischen CK und der Urteilsgüte von Lehrkräften festgestellt werden. Möglicherweise ist für die Einschätzung von Schülerleistungen insbesondere der professionelle Wissensinhalt PCK von Bedeutung (siehe Befunde aus der Domäne Mathematik, z. B. unter 5.7).

8 Ableitung der Forschungsanliegen und Hypothesen

In den vorausgehenden Kapiteln wurde das Konzept der professionellen Kompetenz von Lehrkräften vorgestellt. Verdichtet hat sich die Argumentation darauf, dass professionelle Wissensinhalte von Lehrkräften – und insbesondere die Wissensdomäne PCK – einen bedeutsamen Einfluss auf die Unterrichtsqualität und die Schülerleistung nehmen.

Ergebnisse des systematischen Bildungsmonitorings in Deutschland weisen – trotz Bemühungen des Bildungssystems im Sinne von Fördermaßnahmen für Schüler – darauf hin, dass nach wie vor von einem verheerend hohen Anteil an Schülern auszugehen ist, der nicht über ein ausreichendes Maß an Lesekompetenz verfügt, was einer Beschulung in höheren Klassenstufen im Wege steht.

Da Studien in den naturwissenschaftlichen Domänen positive Zusammenhänge zwischen PCK und der Unterrichtsqualität zeigen (z. B. Kunter, Kleickmann et al., 2011), wird von einer Übertragbarkeit dieses Zusammenhangs auf die geisteswissenschaftlichen Unterrichtsfächer ausgegangen. Daher kann vermutet werden, dass der professionelle Wissensinhalt PCK gleichermaßen bedeutsam für die Unterrichtsqualität im Leseunterricht ist. Demnach sollten (angehende) Deutschlehrkräfte über ein ausreichendes Maß an PCK verfügen, um die Lesekompetenz der eignen Schüler systematisch zu verbessern und aufzubauen. Dies kann durch die Förderung und umfassende Schulung angehender Deutschlehrkräfte hinsichtlich lesedidaktischer Fähigkeiten – bereits während der Lehramtsausbildung – erreicht werden.

Um dies zu ermöglichen, sind empirisch gesicherte Befunde zu Struktur und Entwicklung des lesedidaktischen Wissens von Lehramtsstudierenden und Lehrkräften notwendig, die bislang nicht umfassend vorliegen. Um Erkenntnisse zu Struktur und Entwicklung des lesedidaktischen Wissens zu gewinnen, ist ein standardisiertes, reliables und valides Testinstrument erforderlich (König, 2015). Das Ziel dieser Dissertation ist es, ein ebensolches Testinstrument für den Bereich der Lesedidaktik zu entwickeln.

Das lesedidaktische Wissen wird als professioneller Wissensinhalt von Lehrkräften mit impliziten und expliziten Anteilen aufgefasst (vgl. Kapitel 5.4). Diese Studie zielt darauf ab, auch die impliziten Anteile des lesedidaktischen Wissens zu erfassen. Dies stellt nämlich eine Ergänzung und Erweiterung zu bereits etablierten Testinstrumenten zur Erfassung professioneller Wissensinhalte im Deutsch- bzw. Leseunterricht dar, die vorrangig explizite Bereiche des lesedidaktischen Wissens erfassen (z. B. Bremerich-Vos & Dämmer, 2013).

Um auch implizite Anteile des lesedidaktischen Wissens zu erfassen, eignet sich in besonderem Maße das Testformat des Vignettentests. Durch den advokatorischen Ansatz werden professionelle Wissensinhalte von Lehrkräften handlungsnah abgefragt. Ferner verspricht die Verwendung von Unterrichtsvignetten eine kontextualisierte Erfassung von professionellen Wissensinhalten, wodurch anforderungsnähere Variablen für die Vorhersage des Unterrichtsverhaltens von Lehrkräften geliefert werden (Neuweg, 2015).

Da auch implizite Anteile des lesedidaktischen Wissens erfasst werden sollen, wird in dieser Studie ein geschlossenes Antwortformat verwendet. So soll versucht werden, dem Charakter von impliziten Wissensbeständen gerecht zu werden: Die Probanden müssen die Bewertungen der Testaufgaben nämlich nicht explizit verbalisieren, sondern zwischen verschiedenen konkurrierenden Handlungsalternativen abwägen. So wird der Annahme Rechnung getragen, dass implizite Wissensanteile nicht eindeutig verbalisierbar sind (siehe Kapitel 5.2).

Daneben führt die Verwendung eines ausschließlich geschlossenen Antwortformats dazu, dass das Testinstrument ökonomisch eingesetzt werden kann. Die Forderung nach Ökonomie von Testinstrumenten, die großflächige Befragungen ermöglichen, wird in der deutschdidaktischen Diskussion zunehmend gestellt (Rupp & Bonholt, 2006).

Im Rahmen dieser Dissertationsschrift werden daher folgende drei Forschungsanliegen bearbeitet:

Forschungsanliegen 1: Entwicklung eines Vignettentests mit geschlossenem Antwortformat zur Erfassung des lesedidaktischen Wissens bei angehenden Lehrkräften

Zur Entwicklung der Testaufgaben für den Vignettentest soll zunächst ein theoretisches Strukturmodell des lesedidaktischen Wissens aus der Literatur abgeleitet werden. Die Testentwicklung wird dann auf diesem theoretischen Strukturmodell beruhen (siehe Kapitel 9 *Testentwicklung*).

Das theoretische Strukturmodell des lesedidaktischen Wissens soll anhand einer konfirmatorischen Faktorenanalyse auf empirische Bewährung untersucht werden. *Die Testaufgaben* sollen durch eine IRT-Skalierung auf psychometrische Eigenschaften wie Itemschwierigkeiten und Reliabilität überprüft werden.

Forschungsanliegen 2: Überprüfung der Validität des Vignettentests

Zur Überprüfung der Validität des Vignettentests soll ein Kontrollgruppenvergleich herangezogen werden. So soll sichergestellt werden, dass das neu entwickelte Testinstrument tatsächlich lesedidaktisches Wissen – in Abgrenzung zu anderen professionellen Wissensinhalten – misst. Als Kontrollgruppen zur Hauptstichprobe der Deutschlehramtsstudierenden eignen sich Studierende benachbarter Studiengänge und zwar insbesondere Studierende fachfremder Lehrämter, Psychologiestudierende und Germanistikstudierende. Es wird erwartet, dass die Deutschstudierenden im Testverfahren besser abschneiden als die Kontrollgruppen, da spezifische Wissensinhalte abgefragt werden, die im Deutschstudium vermittelt werden sollten. Die Studierenden fachfremder Lehrämter und die Psychologiestudierenden sollten schlechter abschneiden: Die Studierenden fachfremder Lehrämter beschäftigen sich zwar mit allgemeindidaktischen Prinzipien des Schulunterrichts, allerdings in anderen Domänen als dem Unterrichtsfach Deutsch. Die Psychologiestudierenden beschäftigen sich innerhalb ihres Studiums u. a. mit Lehr-Lernprozessen sowie motivationalen Aspekten des Lernens, allerdings nicht mit fachdidaktischen Inhalten. Am schlechtesten sollten die Germanistikstudierenden abschneiden, da diese innerhalb ihres Studiums weder mit fachdidaktischen Prinzipien noch mit Lehr-Lernprozessen in Berührung kommen. Aus vorhergehenden Studien liegen bislang keine Ergebnisse zu Kontrollgruppenver-

gleichen bezüglich PCK zwischen Studierenden unterschiedlicher Fachrichtungen vor.

Hypothese 1: Lehramtsstudierende mit dem Fach Deutsch schneiden in dem Testverfahren besser ab als Studierende benachbarter Fächer.

Weiterhin wird erwartet, dass das lesedidaktische Wissen von Deutsch-Studierenden über das Studium hinweg (querschnittlich) ansteigt. Verschiedene Autoren haben mit vergleichbaren Testverfahren im Fachbereich Deutsch (Bremrich-Vos & Dämmer, 2013) und anderen Domänen (Baer et al., 2007) gezeigt, dass ein Zuwachs an PCK über das Studium und das Referendariat hinweg abgebildet werden kann. Ein (querschnittlicher) Zuwachs an lesedidaktischem Wissen über das Deutschlehramtsstudium hinweg könnte als Indikator für die curriculare Validität des Vignettentests gewertet werden.

Hypothese 2: Das lesedidaktische Wissen von Lehramtsstudierenden mit dem Fach Deutsch steigt mit zunehmendem Fachsemester (querschnittlich) an.

Weiterhin wird erwartet, dass gewisse Ausbildungsabschnitte bzw. Lerngelegenheiten des Lehramtsstudiums besonders bedeutsam für die Ausbildung und Entwicklung des lesedidaktischen Wissens sind:

Zunächst wird die eigene Lehrerfahrung von Studierenden, die im Rahmen von begleitenden Semesterpraktika erworben wird, als grundlegend für die Entwicklung und Festigung von PCK angesehen. Durch eigene Lehrerfahrungen sollte der Aufbau von PCK durch die Nutzung von akademischem (Fach-)Wissen in authentischen Praxissituationen gefördert werden (Hollins, 2011). Daneben soll durch praktische Ausbildungsabschnitte im Rahmen des Studiums eine Integration von theoretischem Wissen (CK) und praktischem Können (PCK) erfolgen (Hedtke, 2000).

Obwohl Praktika sowohl von Studierenden als auch von Lehrenden als zentrale Lernkontexte während der Lehramtsausbildung bezeichnet werden, blieb ein eindeutiger empirischer Nachweis über die Wirksamkeit und Nachhaltigkeit anhand eines standardisierten Testinstruments bisher aus (Hascher, 2006). Allerdings ge-

ben die Studierenden in Selbstberichtsdaten einen Anstieg fachlicher Kompetenz durch studienbegleitende Praktika an (Hascher, 2006).

Hypothese 3: Die bereits gesammelte Lehrerfahrung der Deutschstudierenden sagt deren lesedidaktisches Wissen vorher.

Es kann weiterhin angenommen werden, dass neben der Lehrerfahrung auch bestimmte Lerngelegenheiten – Seminare, die spezifische lesedidaktische Vorgehensweisen thematisieren – positiv auf die Entwicklung des lesedidaktischen Wissens wirken sollten. Aus dem Fachbereich Mathematik liegen empirische Evidenzen vor, dass der Umfang an mathematischen Lerngelegenheiten signifikant positiv mit dem fachlichen sowie fachdidaktischen Wissen von angehenden Lehrkräften zusammenhängt (Blömeke et al., 2010).

Hypothese 4: Der Besuch spezifischer lesedidaktischer Seminare sagt das lesedidaktische Wissen der Deutschstudierenden vorher.

Diese drei Einflussfaktoren auf das lesedidaktische Wissen – die Studiendauer, die Lehrerfahrung und der Besuch lesedidaktischer Seminare – lassen sich wie folgt charakterisieren:

Die Studiendauer ist als unspezifischer Prädiktor des lesedidaktischen Wissens zu verstehen, da sie „nur“ angibt, in welchem Studienabschnitt sich die Studierenden befinden.

Die Lehrerfahrung und der Besuch lesedidaktischer Seminare sind als spezifische Prädiktoren des lesedidaktischen Wissens zu verstehen: Insbesondere durch die Nutzung dieser Lerngelegenheiten sollte das lesedidaktische Wissen der Studierenden ansteigen. Die Lehrerfahrung und der Besuch lesedidaktischer Seminare werden daher zu dem Konstrukt *praktische Erfahrung* zusammengefasst.

Die sich anschließende Fragestellung ist nun, ob die Studiendauer sowie die praktische Erfahrung der Studierenden inkrementelle Prädiktoren des lesedidaktischen Wissens sind. Neuweg (2014) nimmt in seinem Modell „Konzepte des ‚Lehrer-

wissens“ an, dass sowohl Wissensbereich 1¹⁰ (Wissen, das im ersten Ausbildungsabschnitt in der Lehramtsausbildung erworben wird) als auch „Lernen durch Erfahrung“¹¹ den Aufbau von PCK beeinflusst (siehe Abbildung 2). Es wird daher angenommen, dass die Studiendauer und die praktische Erfahrung der Studierenden inkrementelle Prädiktoren auf das lesedidaktische Wissen bilden.

Hypothese 5: Die Semesteranzahl und die praktische Erfahrung der Studierenden leisten einen inkrementellen Beitrag zur Vorhersage des lesedidaktischen Wissens.

Forschungsanliegen 3: Einbettung des lesedidaktischen Wissens in den breiteren Kontext der professionellen Kompetenz von Lehrkräften

Das Modell der professionellen Kompetenz von Lehrkräften (siehe Abbildung 1) postuliert sowohl Zusammenhänge zwischen den verschiedenen inhaltlichen Dimensionen des professionellen Wissens (PK, CK und PCK) als auch zwischen weiteren kognitiven und affektiven Teilkompetenzen. Konkrete Befunde, wie diese verschiedenen Kompetenzaspekte im Gesamten miteinander verknüpft sind, liegen allerdings bisher kaum vor. Häufig werden in Forschungsarbeiten (bezüglich einzelner Domänen, z. B. Mathematik) Zusammenhänge der verschiedenen Teilkompetenzen miteinander oder der Einfluss einzelner Teilkompetenzen auf die Schülerleistung untersucht, z. B. zwischen der motivationalen Orientierung der Lehrkräfte und der Schülerleistung (Kunter, 2011). Viele Forschungsarbeiten befassen sich daneben mit Zusammenhängen zwischen den unterschiedlichen Inhaltsbereichen des professionellen Wissens, also zwischen PK, CK und PCK (Kunter, Baumert et al., 2011). Daneben bilden Persönlichkeitseigenschaftsvariablen (z. B. BIG-5) von (angehenden) Lehrkräften eine der (Lern-)Voraussetzungen zum Erwerb professioneller Kompetenzen (Baumert & Kunter, 2006; Kunter, Kleickmann et al., 2011; Mayr, 2014).

Ein wesentliches Anliegen dieser Studie ist daher, den Zusammenhang des lesedidaktischen Wissens mit weiteren Aspekten der professionellen Kompetenz von

¹⁰ Hier operationalisiert über die Semesteranzahl.

¹¹ Hier operationalisiert über den Indikator *praktische Erfahrung*.

Deutschlehramtsstudierenden zu explorieren¹². Hierzu sollen neben dem lesedidaktischen Wissen weitere kognitive und affektive Konstrukte erhoben werden, die einen Zusammenhang zum professionellen Wissen haben sollten.

Daher wurden folgende Konstrukte mit empirisch bewährten Instrumenten erfasst, um die erwarteten Zusammenhänge mit dem lesedidaktischen Wissen zu untersuchen: PK (König & Blömeke, 2010), lesedidaktische Überzeugungen (Behrmann & Souvignier, 2013), das berufsbezogene Selbstkonzept (Retelsdorf, Bauer, Gebauer, Kauper & Möller, 2014a), Berufswahlmotive (Pohlmann & Möller, 2010) und Persönlichkeitseigenschaften (Gerlitz & Schupp, 2005).

Bezüglich PK wird erwartet, dass ein positiver Zusammenhang zum lesedidaktischen Wissen gezeigt werden kann. Da sich PCK sowohl aus pädagogisch-psychologischen als auch aus Fachwissensanteilen speist, ist diese Vermutung theoretisch ableitbar (vgl. Shulman, 1986, 1987). Verschiedene Studien haben von unterschiedlich starken Zusammenhängen zwischen PK und PCK berichtet, z. B. bei Physiklehrkräften ($r = 0.27$, $p \leq 0.001$, Kirschner, 2013), Techniklehrkräften ($r = 0.40$, $p \leq 0.001$, Lin, Tsai, Chai & Lee, 2013) und angehenden Biologielehrkräften ($r = 0.32$, $p \leq 0.001$, Paulick, Großschedl, Harms & Möller, 2016).

Hypothese 6: Das lesedidaktische Wissen (PCK) weist einen positiven Zusammenhang zu PK auf.

Bezüglich lesedidaktischer Überzeugungen soll untersucht werden, ob eine konstruktivistische lesedidaktische Überzeugung einen positiven Zusammenhang zum lesedidaktischen Wissen aufweist. Wie unter 4.2 dargestellt spielen Überzeugungen von Lehrkräften eine bedeutsame Rolle für das Unterrichtsgeschehen und schließlich auch für die Schülerleistung (Voss et al., 2011). In der Literatur gibt es Hinweise darauf, dass eine konstruktivistische Sichtweise auf Lehr-Lern-Prozesse (Savery & Duffy, 1995) im Gegensatz zu einer direkt-transmissiven Sichtweise (Resnick & Hall, 1998) mit einer besseren Schülerleistung einhergeht (Souvignier & Mokhlesgerami, 2005). Daher ist zielführend zu untersuchen, inwiefern diese

¹² Selbstverständlich kann hier aus ökonomischen Gründen sowie einem Mangel an geeigneten Instrumenten nur eine Auswahl an interessierenden Konstrukten untersucht werden.

für den Leseunterricht spezifischen Überzeugungen mit dem lesedidaktischen Wissen in Zusammenhang stehen. In Bezug auf Mathematiklehrkräfte konnte ein Zusammenhang zwischen einer konstruktivistischen Überzeugung und PCK festgestellt werden ($r = 0.32$, $p \leq 0.05$, Kunter et al., 2013).

Hypothese 7: Das lesedidaktische Wissen weist einen positiven Zusammenhang zu einer konstruktivistischen lesedidaktischen Überzeugung und einen negativen Zusammenhang zu einer direkt-transmissiven lesedidaktischen Überzeugung auf.

Bezüglich des berufsbezogenen Selbstkonzeptes wird erwartet, dass Lehramtsstudierende mit einem stärker ausgeprägten berufsbezogenen Selbstkonzept eine bessere Testleistung erzielen als Studierende mit einem weniger stark ausgeprägten berufsbezogenen Selbstkonzept. Die Annahme, dass das Selbstkonzept prädiktiv für (akademische) Leistungen innerhalb einer Domäne ist (Möller & Trautwein, 2015), wurde in verschiedenen empirischen Studien bestätigt (Marsh & Martin, 2011). Bezogen auf Lehramtsstudierende konnte gezeigt werden, dass die Studienleistungen von Lehramtsstudierenden statistisch bedeutsame Zusammenhänge zu Bereichen des selbstberichteten berufsbezogenen Selbstkonzeptes aufweisen (Retelsdorf et al., 2014b). Ferner kann die Skala zum Selbstkonzeptaspekt Agency den objektiven¹³ sowie subjektiven Berufserfolg von Lehrkräften prädictieren (Abele, 2011).

Hypothese 8: Ein stärker ausgeprägtes berufsbezogenes Selbstkonzept geht mit einer besseren Testleistung einher.

Bezüglich der Berufswahlmotive wird untersucht, inwiefern intrinsische und extrinsische Aspekte der Berufswahlmotivation Zusammenhänge mit dem lesedidaktischen Wissen aufweisen. Im schulischen Kontext konnte gezeigt werden, dass eine intrinsische Motivation der Schüler einen positiven Zusammenhang zu Schulnoten und der Leistung in standardisierten Testverfahren aufweist, extrinsi-

¹³ Zur Operationalisierung von objektivem und subjektivem Berufserfolg bei Abele (2011) siehe Kapitel 4.2.

sche Motivation dagegen einen negativen (Lepper, Corpus & Iyengar, 2005). Bezüglich Lehramtsstudierender liegen empirische Evidenzen vor, dass intrinsische Motivationsaspekte mit einer höheren durchschnittlichen akademischen Leistung im Lehramtsstudium sowie einem stärker ausgeprägten Wunsch Lehrer zu werden einhergehen (Uyulgan & Akkuzu, 2014). Daneben führt intrinsische Motivation bezüglich eines gewissen Lerngegenstandes zu einem vertieften Engagement und mehr Ausdauer beim Lernen sowie einem verbesserten konzeptuellen Lernen (Vansteenkiste, Leny & Deci, 2006). Da in der Literatur eine konsistente Befundlage zum positiven Einfluss von intrinsischer Motivation auf akademische Leistungen sowie verschiedene Lernaktivitäten vorliegt, werden für intrinsische Aspekte der Berufswahlmotivation daher positive Zusammenhänge mit dem lesedidaktischen Wissen angenommen.

Hypothese 9 (explorativ): Das lesedidaktische Wissen weist einen positiven Zusammenhang zu intrinsischen Aspekten der Berufswahlmotivation auf.

In Bezug auf die BIG-5 wird erwartet, dass die Merkmale Extraversion und Gewissenhaftigkeit einen positiven Zusammenhang mit dem lesedidaktischen Wissen aufweisen. Mayr (2014) und Hanfstingel und Mayr (2007) haben im Rahmen einer Meta-Analyse Zusammenhänge zwischen Persönlichkeitseigenschaften von (angehenden) Lehrkräften und deren Bewährung im Lehrerberuf untersucht. In dieser Meta-Analyse konnte gezeigt werden, dass Extraversion und Gewissenhaftigkeit einen statistisch bedeutsamen Zusammenhang zur pädagogischen Handlungskompetenz im Praktikum¹⁴ aufweisen. Es wird angenommen, dass der pädagogischen Handlungskompetenz im Praktikum vorrangig der professionelle Wissensinhalt PCK zugrunde liegt. Daher wird analog zu diesen Befunden davon ausgegangen, dass Extraversion und Gewissenhaftigkeit mit dem lesedidaktischen Wissen der Studierenden in Zusammenhang stehen sollten.

¹⁴ Der Begriff „pädagogische Handlungskompetenz im Praktikum“ fasst in der Meta-Analyse von Hanfstingel und Mayr (2007) die Einschätzungen unterschiedlicher Teilkompetenzen (z.B. im Erziehen, Unterrichten) sowie Angaben zum pädagogisch-didaktischen Handeln der angehenden Lehrkräfte zusammen.

Hypothese 10: Das lesedidaktische Wissen weist einen positiven Zusammenhang zu den Persönlichkeitsmerkmalen Extraversion und Gewissenhaftigkeit auf.

9 Testentwicklung

Die Ableitung eines theoretischen Strukturmodells des lesedidaktischen Wissens und die darauf aufbauende Testentwicklung bilden das Herzstück dieser Studie. Im folgenden Kapitel werden die theoretischen Grundlagen des Vignettentests ausführlich beschrieben. Anschließend werden die Ergebnisse der mehrstufigen inhaltlichen Validierung der Testaufgaben durch qualitative und quantitative Expertenbefragungen vorgestellt. Zuletzt wird erläutert, wie die Lösungsmuster für die Unterrichtsvignetten anhand einer Expertennorm erstellt wurden.

9.1 Ableitung eines theoretischen Modells des lesedidaktischen Wissens für die Sekundarstufe I

Wie oben dargelegt, liegen bislang nur eingeschränkte empirische Ergebnisse zu professionellen Wissensinhalten von Deutschlehramtsstudierenden vor (Bremrich-Vos & Dämmer, 2013). Die erste Aufgabe in der Testentwicklung bestand daher darin, ein theoretisches Modell des lesedidaktischen Wissens aus der lesedidaktischen, allgemein fachdidaktischen und kognitionspsychologischen Literatur abzuleiten, auf dessen Grundlage die Testentwicklung fußen konnte¹⁵. Die vorliegende Studie befasst sich mit dem lesedidaktischen Wissen von Lehramtsstudierenden für das Unterrichten in der Sekundarstufe I. Hier geht es im Gegensatz zur Primarstufe, wo der Erstleseunterricht zentral ist, hauptsächlich um den Lese- und Literaturunterricht.

Bei der Ableitung des Modells des lesedidaktischen Wissens wurden verschiedene theoretische Überlegungen berücksichtigt. Shulman (1986) hebt zwei Facetten von PCK besonders hervor: (1) Das Wissen über Erklären und Darstellen und (2) das Wissen über fachbezogene Schülerkognitionen. Diese Annahme zweier zentraler Kernbereiche von PCK hat eine Vielzahl von nachfolgenden Autoren unterstützt und bestätigt (siehe Kapitel 5.6).

¹⁵ Das theoretische Strukturmodell diene als Arbeitsmodell zur Erstellung der Testaufgaben und ist nicht als Strukturgleichungsmodell im eigentlichen Sinne zu verstehen.

Aus dem Bereich der deutschdidaktischen Forschung liegen theoretische Überlegungen zu Modellen der Deutschdidaktik vor (Helmers 1976; Ossner 2006). Diese Modelle befassen sich global mit den Anforderungen an Lehrkräfte im Deutschunterricht und sind aus diesem Grund sehr komplex. Eine empirische Prüfung konnte daher bislang nicht stattfinden. Überlegungen aus der deutschdidaktischen Forschung zu gelingender Lesedidaktik werden insbesondere bei der inhaltlichen Ausrichtung der Unterrichtsvignetten herangezogen.

Seitens der kognitionspsychologischen Forschung zum Lesen haben Kintsch und van Dijk (1978) ein empirisch bewährtes Prozessmodell des Leseverstehens vorgelegt (siehe Kapitel 6.2). In diesem Modell postulieren die Autoren ein System von verschiedenen mentalen Operationen auf verschiedenen Hierarchieebenen, die das Leseverstehen ermöglichen. Diese konkreten kognitionspsychologischen Abläufe beim Leseprozess sind allerdings nicht Gegenstand des entwickelten Testinstruments, welches das lesedidaktische Wissen abbilden soll. Konkretes Wissen über den Leseprozess ist dem Wissensbereich CK zuzuordnen.

Zur Entwicklung des Strukturmodells wurden außerdem die Standards in der Lehrerbildung herangezogen (Kultusministerkonferenz, 2014). Die Kultusministerkonferenz nennt darin vier professionelle Kompetenzbereiche, die Lehramtsstudierende während ihrer Ausbildung erwerben sollten: Unterrichten, Erziehen, Bewerten und Innovieren. Betrachtet man den professionellen Wissensinhalt PCK, so ist dieser inhaltlich insbesondere den Kompetenzbereichen Unterrichten und Bewerten zuzuordnen.

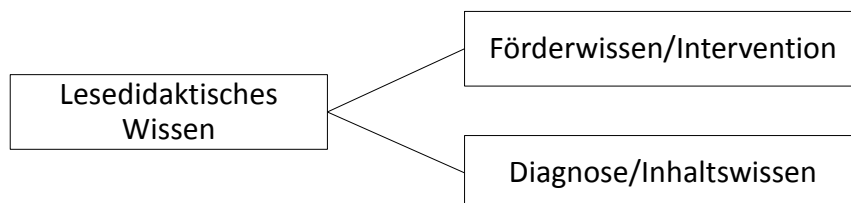


Abbildung 5. Theoretisches Strukturmodell des lesedidaktischen Wissens.

Die oben dargestellten Vorarbeiten wurden in ein zweidimensionales Modell des lesedidaktischen Wissens überführt. Zweidimensional, da die überwiegende

Mehrheit von einschlägigen Forschungsarbeiten zwei inhaltliche Facetten für den professionellen Wissensbereich PCK annimmt (siehe Kapitel 5.6).

Die erste Facette „Förderwissen/Intervention“ meint das aktive fachdidaktische Handeln einer Lehrperson im Leseunterricht; hier ist z. B. die Fähigkeit der Lehrperson gemeint, adäquat auf Verständnisschwierigkeiten beim Lesen eines Textes zu reagieren und entsprechende Hilfestellungen geben zu können. Hier fließen die Facette „Wissen über Erklären und Darstellen“ von Shulman (1986,1987) und der Kompetenzbereich „Unterrichten“ der Kultusministerkonferenz (2014) mit ein.

Unter 7.1.1 bis 7.1.3 wurden fachdidaktische Vorgehensweisen vorgestellt, über die (angehende) Deutschlehrkräfte verfügen sollten, um die Lesekompetenz von Schülern im Rahmen des schulischen Unterrichts gewinnbringend zu fördern (Artelt & Dörfler, 2010; Rosebrock & Nix, 2014; Streblow, 2004). Als besonders effektiv haben sich Verfahren zur Steigerung der Leseflüssigkeit (Fuchs, Fuchs, Hosp & Jenkins, 2001; Rasinski, 2005), der Lesemotivation (Deci & Ryan, 1985; Mucherah & Ambrose-Stahl, 2014) sowie zum vermehrten Einsatz von (meta-)kognitiven Lesestrategien erwiesen (Carretti, Caldarola, Tencati & Cornoldi, 2014; Lenhard, Baier, Endlich, Schneider & Hoffmann, 2013).

Die zweite Facette „Diagnose/Inhaltswissen“ meint das inhaltsbezogene fachdidaktische Wissen von Lehrkräften. Ein Beispiel hierfür wäre, ob eine Lehrperson angemessen schwierige Leseaufgaben für eine bestimmte Klasse auswählen oder Lese- und Verstehensleistungen von Schülern treffend einschätzen kann. Hier fließen die Facette „Wissen über fachbezogene Schülerkognitionen“ von Shulman (1986, 1987) und der Kompetenzbereich „Bewerten“ der Kultusministerkonferenz (2014) mit ein.

Unter 7.1.4 wurde erläutert über welche Wissensbestände (angehende) Deutschlehrkräfte verfügen sollten, um die Leseleistung der Schüler korrekt und zuverlässig diagnostizieren zu können. Insbesondere Kenntnisse über die mentalen Abläufe während des Leseprozesses befähigen Lehrkräfte dazu zuverlässig festzustellen, in welchen Teilbereichen des Leseprozesses Einschränkungen vorhanden sind. Daneben ist ein fundiertes Wissen über Aufgabentypen sowie schwierigkeitsgenerierende Textmerkmale vonnöten.

Zusammenfassend kann festgestellt werden, dass das theoretische Strukturmodell des lesedidaktischen Wissens aus verschiedenen Überlegungen zum Professionalisierungsprozess von (Deutsch-)Lehramtsstudierenden abgeleitet ist. Das Modell postuliert zwei inhaltlich voneinander trennbare inhaltliche Wissensfacetten, die erwartungsgemäß einen mittleren bis hohen Zusammenhang aufweisen (Park & Chen, 2012).

9.2 Entwicklung der Testaufgaben

Auf der Grundlage des theoretischen Strukturmodells des lesedidaktischen Wissens für die Sekundarstufe I wurden Vignetten entwickelt, die sich auf Problemstellungen im Leseunterricht beziehen.

Im Entwicklungsprozess wurde bewusst darauf verzichtet, die Unterrichtsvignetten an konkreten Inhalten der Curricula einzelner Hochschulen zu orientieren. Diese Studie intendiert nicht das Erlernen spezifischer curricularer Inhalte abzu prüfen, sondern fokussiert vielmehr allgemeine lesedidaktische Wissensinhalte, die losgelöst von hochschulspezifischen Curricula erworben werden sollten. Es ist nämlich wünschenswert, dass (angehende) Deutschlehrkräfte unabhängig von der Hochschulzugehörigkeit über tiefgreifendes Wissen bezüglich des Leseprozesses sowie adäquate fachdidaktische Vorgehensweisen zur Steigerung der Lesekompetenz der eigenen Schüler verfügen. Daher wurden die Unterrichtssituationen angelehnt an die – hochschulübergreifend bekannten – Forschungsbefunde zur Lesedidaktik entwickelt.

Die Vignetten sind folgendermaßen aufgebaut (siehe Abbildung 6): Jede Vignette hat eine kurze Überschrift, die knapp vorwegnimmt, welches Thema die Unterrichtssituation haben wird. Dann folgt die kurze Beschreibung der Unterrichtssituation aus dem Leseunterricht (sog. *Vignettenstamm*). In dieser kurzen Beschreibung werden die Probanden über den bisherigen Ablauf der Unterrichtsstunde bzw. der Unterrichtseinheit informiert. Ferner werden die Klassenstufe (Klassenstufen 5 bis 10) und die Schulart (Hauptschule, Werkrealschule, Gesamtschule oder Realschule) genannt.

Es folgt eine fachdidaktisch relevante Fragestellung, die sich darauf bezieht, wie die fiktive Lehrperson in der gegebenen Unterrichtssituation fachdidaktisch sinnvoll reagieren bzw. den Unterricht angemessen fortführen könnte.

Zu der fachdidaktisch relevanten Fragestellung werden Aussagen (Items) präsentiert, die auf einer sechsstufigen Likert-Skala¹⁶ eingeschätzt werden sollen.

Überprüfen des Textverständnisses

Die Schüler(innen) einer 8. Hauptschulklasse haben eine Kurzgeschichte mit dem Titel „Beste Freunde, egal was passiert“ gelesen. Der Text ist gut gegliedert und besteht überwiegend aus kurzen Sätzen. Zuvor hat sich die Klasse auf das Thema eingestimmt, indem sie in einer Mindmap erarbeitet hat, was den Schüler(inne)n in einer Freundschaft wichtig ist.

Der Lehrer möchte nun überprüfen, ob die Schüler(innen) den Inhalt des Textes erfasst haben.

Wie könnte der Lehrer aus fachdidaktischer Sicht überprüfen, ob die Schüler(innen) den Inhalt des Textes erfasst haben?

Machen Sie in jeder Zeile ein Kreuz.

		Trifft gar nicht zu					Trifft voll und ganz zu
		1	2	3	4	5	6
a)	Der Lehrer sollte die Schüler(innen) fragen, ob sie den Text verstanden haben.	[]	[]	[]	[]	[]	[]
b)	Die Schüler(innen) sollten in Einzelarbeit Fragen an den Text formulieren.	[]	[]	[]	[]	[]	[]
c)	Der Lehrer sollte, während die Klasse liest, herumgehen und einzelnen Schüler(inne)n Fragen zum Text stellen.	[]	[]	[]	[]	[]	[]
d)	Die Schüler(innen) sollten den Inhalt der Kurzgeschichte auf die zuvor erarbeitete Mindmap beziehen.	[]	[]	[]	[]	[]	[]
e)	Die Schüler(innen) sollten eine Fortsetzung der Kurzgeschichte schreiben.	[]	[]	[]	[]	[]	[]

Abbildung 6. Beispielvignette aus dem Testinstrument zur Facette „Diagnose/Inhaltswissen“.

Insgesamt wurden 31 Vignetten mit 8 bis 15 Items generiert. Wie oben dargestellt wurden die Erstellung der Vignetten thematisch an den verschiedenen Forschungsbereichen zur Lesekompetenz und Lesedidaktik orientiert. In der unten stehenden Tabelle befindet sich ein Überblick über die Zuordnung der Vignetten zu den beiden Modellfacetten (siehe Tabelle 2). Es wird deutlich, dass nicht zu jeder Modellfacette Vignetten aus jedem Forschungsbereich entwickelt wurden. Ur-

¹⁶ Die verwendete sechsstufige Skala ist genau genommen als *Likert-ähnlich* zu bezeichnen, da eine Äquidistanz zwischen den einzelnen Skalenpunkten nicht eindeutig angenommen werden kann.

sächlich hierfür ist, dass sich die verschiedenen lesedidaktischen Situationen unterschiedlich gut zu den Modellfacetten zuordnen lassen.

Tabelle 2. Zuordnung der Vignetten zu Modellfacette und lesedidaktischem Forschungsbereich

Forschungsbereich	Modellfacette	
	Förderwissen/Intervention	Diagnose/Inhaltswissen
	Lesestrategien	Informationen im Text lokalisieren (6)
		Einen Sachtext mündlich präsentieren (10)
		Gruppenarbeit zum Thema „Klimawandel“ (9)
		Die Lesestrategien Mindmap (10)
		Die Fragen sind voll schwer! (13)
		Einen anspruchsvollen Text lesen (15)
	Lesemotivation	Ein Jugendbuch lesen (11)
		Vorlesen vor der Klasse (13)
		„Ich bin noch nicht fertig mit Lesen“
		„Ich kapiere die Bücher eh nie...“ (9)
		Lautlesen will geübt sein (12)
		Die Bücherkiste (11)
	Leseflüssigkeit	Lesetrainer und Lesesportler (7)
		Wie kann die Leseflüssigkeit gesteigert werden? (8)
		Wählt euch ein Buch selbst aus! (12)
	Textverstehen	Welches Sprichwort passt? (11)
		„Mein Freizeitverhalten passt nicht zur Graphik!“ (10)
		Hat Sonja das Zitat verstanden? (11)
		„Es geht um Stress bei der Arbeit“ (11)
		Ich verstehe den Satz nicht! (11)
		Emil und die Detektive (10)

Textschwierigkeit	Die Fabel „Der Löwe und der Bär“ (9)
	Was machen Jugendliche in ihrer Freizeit? (8)
	Der Text „Die Kraft der Siedler“ (11)
	„Geier“ (12)
	Überprüfen des Textverständnisses (13)
Aufgabenschwierigkeit	Wie kann man einen Text vereinfachen? (16)
	Eine Aufgabe zur Fabel „Der Löwe und der Bär“ (9)
	Eine Graphik beschreiben (6)
	Eine Aufgabe zum Text „Die Kraft der Siedler“ (11)
	Um was geht es hier? (9)

Anmerkung. Die Anzahl der entwickelten Items sind in Klammern hinter den Vignetten-titeln abzulesen.

9.3 Inhaltliche Validierung der Vignetten und Erstellung des Lösungsmusters

Sowohl die Vignetten(-stämme) als auch die Items innerhalb der Vignetten wurden einem mehrstufigen Prozess der inhaltlichen Validierung unterzogen. Die aufeinanderfolgenden Validierungsschritte sind in Abbildung 7 übersichtlich dargestellt. Eine ausführliche Beschreibung der Validierungsschritte und deren Ergebnisse finden sich in den nachfolgenden Abschnitten dieses Kapitels.

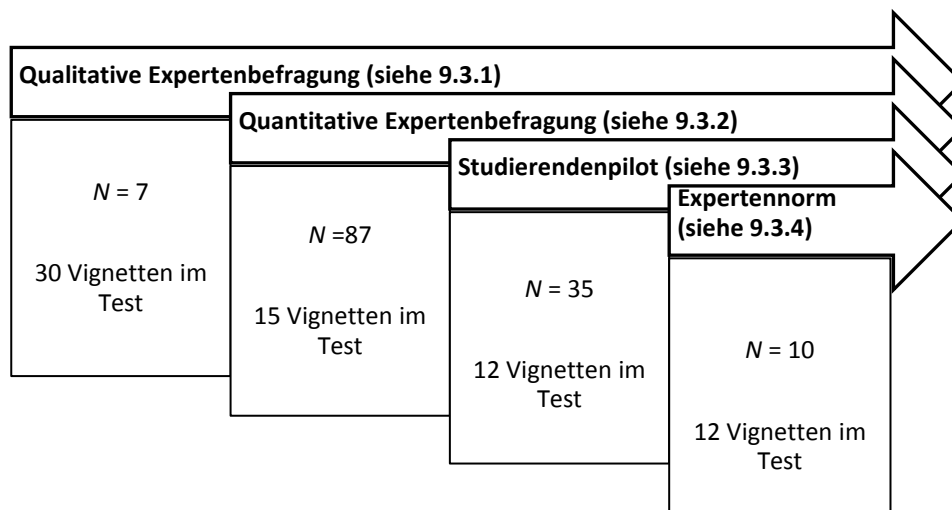


Abbildung 7. Darstellung des Prozesses der inhaltlichen Validierung der Testaufgaben.

Neben der inhaltlichen Validierung des Vignettentests sollten durch dieses dezierte Vorgehen diejenigen Vignetten und Items für das finale Testinstrument ausgewählt werden, die laut Expertenmeinung am besten für die Erfassung des lesedidaktischen Wissens von angehenden Lehrkräften geeignet sind.

9.3.1 Qualitative Expertenbefragung

Einen ersten Schritt zur inhaltlichen Validierung der Vignetten stellte eine qualitative Expertenbefragung dar. Das Ziel dieses ersten Validierungsschrittes war es, die Vignetten hinsichtlich Eindeutigkeit, Verständlichkeit und fachdidaktischer Relevanz zu optimieren.

Hierzu wurden Experten ($N = 7$, Alter: wurde im Interview nicht abgefragt, 72 % weiblich) aus dem Bereich der Lesedidaktik aus den Staatlichen Seminaren für Didaktik und Lehrerbildung aus Ludwigsburg und Karlsruhe in einem halbstandardisierten qualitativen Interview befragt. Es wurden jeweils unterschiedliche Vignetten vorgelegt; dabei wurde jede Vignette von mindestens vier Experten begutachtet. Die Experten wurden gebeten, die fachdidaktische Relevanz der einzelnen Vignetten einzuschätzen und die Eindeutigkeit der Unterrichtssituationen durch kritische Reflexion zu schärfen. Zwischen den Interviewterminen wurden die Änderungsvorschläge in die Vignetten eingearbeitet. Zusätzlich wurden die

Experten gebeten, die Items hinsichtlich ihrer Eignung zu beurteilen. Vorschläge der Experten für zusätzliche Items wurden in den Test eingearbeitet.

Nach der qualitativen Expertenbefragung wurde eine Vignette aus dem Testinstrument entfernt, sodass 30 Vignetten in den nächsten inhaltlichen Validierungsschritt, eine quantitative Expertenbefragung, übergangen.

9.3.2 Quantitative Expertenbefragung

In einem zweiten inhaltlichen Validierungsschritt wurde eine umfangreiche quantitative Expertenbefragung ($N = 87$; Alter: $M = 46.54$ Jahre, $SD = 10.59$, 64 % weiblich) mit Experten aus dem Bereich der Lesedidaktik aus unterschiedlichen Institutionen durchgeführt (siehe Abbildung 8).

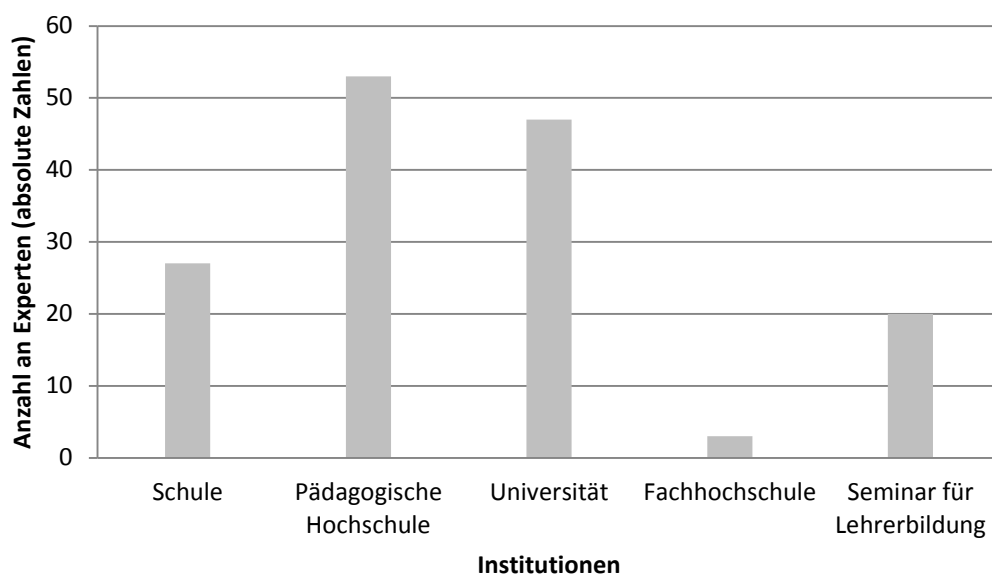


Abbildung 8. Übersicht über die Institutionen, an denen die Experten tätig sind. Mehrfachnennungen sind möglich.

Das Ziel dieser quantitativen Expertenbefragung war es, wenig(er) geeignete Vignetten und Items aus dem Test zu entfernen, um diesen valider und ökonomischer zu gestalten. Die Auswahl der Experten erfolgte anhand einer Recherche an Pädagogischen Hochschulen in Baden-Württemberg und in der Schweiz, Universitäten in Deutschland und den Kooperationspartnern des FuN-Kollegs „Effektive Kom-

petenzdiagnose in der Lehrerbildung“ (EKoL)¹⁷, den Staatlichen Seminaren für Lehrerbildung Ludwigsburg und Karlsruhe; als Experten galten Personen, die sich in ihrer Haupttätigkeit wissenschaftlich oder praktisch mit Lesedidaktik beschäftigen. Insgesamt wurden 398 Personen identifiziert, die als Experte in Frage kamen. An dieser Stelle innerhalb der Testentwicklung war es beabsichtigt, eine möglichst breite Befragung von Experten aus unterschiedlichen Tätigkeitsbereichen durchzuführen, um verschiedene Denkart der Lesedidaktik mit in die Testentwicklung einbeziehen zu können.

Die Experten wurden per E-Mail zur Teilnahme an der Expertenbefragung eingeladen (Rücklaufquote: 87 von 398 kontaktierten Experten, 21.9 %). Es wurde dabei mit dem Online-Tool SoSci-Survey gearbeitet. Die Aufgabe der Experten war es, die Vignetten zu „bearbeiten“, d. h. die Experten füllten den Vignettentest wie Probanden aus. Außerdem konnten die Experten zu jeder Vignette Kritik oder Änderungsvorschläge in einem offenen Textfeld anbringen.

Die 30 Vignetten waren in zwei Testteile aufgeteilt, um die zeitliche Belastung durch das Bearbeiten des Vignettentests für die Experten im Rahmen zu halten. Jede Vignette wurde daher von ungefähr 40 Experten beurteilt (siehe Tabelle 3). Des Weiteren wurden die Vignetten sowie die Items permutiert präsentiert, um Ermüdungs- und/oder Reihenfolgeeffekte kontrollieren zu können.

Bei der Reduktion von Vignetten wurden diejenigen Vignetten aus dem Test ausgeschlossen, die als am wenigsten eindeutig beschrieben sowie im geringsten Ausmaß als fachdidaktisch relevant eingeschätzt wurden; bei der Reduktion von Items wurde darauf geachtet, dass sich die Experten möglichst einig bei der Beurteilung eines Items sind.

Die Reduktion der Items wurde an Tepner und Dollny (2014) orientiert. Diese schlagen ein mehrstufiges Vorgehen zur Auswahl der am besten geeigneten Items bei Unterrichtsvignetten mit geschlossenem Antwortformat vor. Das Ziel dieses mehrstufigen Vorgehens ist diejenigen Items auszuwählen, die die Experten am einheitlichsten bzw. eindeutigsten eingeschätzt haben. Es wurde folgendermaßen vorgegangen:

¹⁷ Diese Dissertationsschrift entstand im Rahmen eines Teilprojektes des Forschungs- und Nachwuchskollegs EKoL („Effektive Kompetenzdiagnose in der Lehrerbildung“) an den Pädagogischen Hochschulen Heidelberg und Ludwigsburg.

1. Es wurden Items entfernt, wenn mehr als 25 % der Experten eine Abweichung $|x| > 2$ vom Modalwert der Antwortverteilung aufwiesen.
2. Es wurden Items ohne ganzzahligen Modalwert entfernt.
3. Es wurde überprüft, ob die verbliebenen Items die gesamte Skalenbreite abbildeten (innerhalb jeder Vignette sollten Items mit niedrigen, mittleren und hohen Ausprägungen enthalten sein).
4. Es wurde die Standardabweichung des Modalwert überprüft: Wenn zwei Items mit der gleichen Ausprägung im Test verbleiben würden, wurde das Item mit der höheren Standardabweichung vom Modalwert entfernt. Eine geringere Standardabweichung indiziert eine geringere Streuung der Itemeinschätzung und somit eine höhere Einigkeit der Experten.

Nach dieser Analyse verblieben 27 Vignetten à vier bis sechs Items im Test. Um die Vignetten für die Pilotstudie weiter zu reduzieren und damit zu ökonomisieren, wurden die Bewertungen der Unterrichtssituationen (abgefragt über drei Items: Eindeutigkeit, fachdidaktische Relevanz, Nähe zum Unterrichtsalltag, sechsstufige Likert-Skala, 1 = *Trifft gar nicht zu*, 6 = *Trifft voll und ganz zu*) herangezogen. Nach dieser Analyse verblieben 15 Vignetten à vier bis sechs Items im Testinstrument.

Tabelle 3. Einschätzung der Vignetten in der quantitativen Expertenbefragung

Einschätzung der Experten (1 = Trifft gar nicht zu, 6 = Trifft voll und ganz zu)									
	Die Situation ist fachdidaktisch relevant.			Die Situation ist nahe am Unterrichtsalldag.			Die Situation ist eindeutig beschrieben.		
Vignettenkurztitel	N	M	SD	N	M	SD	N	M	SD
Leseeflüssigkeit steigern	40	5.45	0.99	40	5.20	1.18	39	5.28	1.07
Buch auswählen	42	4.95	1.29	42	4.79	1.30	41	5.37	0.94
Lesetrainer und Lesesportler	34	5.32	0.77	34	4.50	1.24	34	5.21	1.04
Emil und die Detektive	37	5.43	0.93	37	5.41	0.93	36	5.42	0.84
Leseverstehen Geier	40	5.40	1.08	40	5.18	1.28	40	5.15	1.21
Ich verstehe den Satz nicht	44	5.30	0.82	44	5.05	1.03	42	5.33	1.14
Ein Jugendbuch lesen	42	5.60	0.91	42	5.50	0.94	42	5.14	1.26
Die Bücherkiste	42	5.60	0.59	42	5.24	1.10	42	5.38	1.03
Ich kapier' die Bücher eh nie	33	5.24	1.09	33	5.30	0.81	33	5.52	0.76
Die Fragen sind voll schwer	34	5.59	0.66	34	5.65	0.54	34	5.47	0.83
Lesestrategien Mindmap	33	5.58	0.71	34	5.62	0.60	33	5.42	0.94
Gruppenarbeit Klimawandel	41	5.61	0.80	41	5.46	0.95	41	5.12	1.29
Kurzgeschichte Freundschaft	42	5.55	0.71	41	5.46	0.71	41	5.02	1.17
Textschwierigkeit Diagramm	34	5.65	0.49	34	5.24	0.92	34	5.03	1.22
Text vereinfachen	33	5.48	0.83	33	5.12	1.14	33	5.30	0.98

Anmerkung. Es werden nur die 15 für die Pilotstudie ausgewählten Vignetten dargestellt.

9.3.3 Pilotierung

Die 15 verbliebenen Vignetten wurden in einer Studierendenstichprobe pilotiert ($N = 35$, Alter: $M = 23.97$, $SD = 3.03$, 91.4 % weiblich). Das Ziel der Pilotstudie war es, die Durchführung der Befragung zu testen. Des Weiteren sollte die Akzeptanz („Wie sinnvoll finden Sie die Beantwortung der Testfragen, um lesedidaktisches Wissen von Studierenden zu erfassen?“) und die empfundene Belastung der Studierenden („Wie beanspruchend war die Beantwortung der Testfragen für Sie?“) überprüft werden. Die Akzeptanz und die empfundene Belastung schätzten die Studierenden auf sechsstufigen Likert-Skalen ein (1 = *Überhaupt nicht sinnvoll* bzw. *Überhaupt nicht beanspruchend*, 6 = *Sehr sinnvoll* bzw. *Sehr beanspruchend*).

Außerdem sollten die Studierenden die Verständlichkeit („Wie verständlich ist die Situationsbeschreibung für Sie?“) der einzelnen Vignetten auf einer sechsstufigen Skala einschätzen (1 = *Überhaupt nicht verständlich*, 6 = *Sehr verständlich*). Die Befragung der Studierenden erfolgte im Rahmen eines Seminars und dauerte ca. 90 Minuten.

Tabelle 4. Beurteilung der Vignetten im Studierendenpilot

	Einschätzungen der Studierenden (1 = geringste Ausprägung, 6 = höchste Ausprägung)		
	<i>M</i>	<i>SD</i>	<i>N</i>
Beanspruchung bei Testbearbeitung	3.17	0.86	35
Akzeptanz der Testaufgaben	4.43	1.07	35
Verständlichkeit „Emil und die Detektive“	5.75	0.44	32
Verständlichkeit „Jugendliche“	5.24	0.94	33
Verständlichkeit „Lautlesen will geübt sein“	5.82	0.39	34
Verständlichkeit „Ich kapiere die Bücher eh nie“	5.35	1.01	34
Verständlichkeit „Lesestrategien“	5.61	0.79	33
Verständlichkeit „Stress bei der Arbeit“	5.62	0.60	34
Verständlichkeit „Bücherkiste“	5.79	0.48	34
Verständlichkeit „Klimawandel“	5.52	0.80	33
Verständlichkeit „Lesetandem“	5.65	0.69	34
Verständlichkeit „Ich verstehe den Satz nicht“	5.56	0.71	34
Verständlichkeit „Textverständnis überprüfen“	5.64	0.65	33
Verständlichkeit „Text vereinfachen“	5.50	0.86	34
Verständlichkeit „Lese Flüssigkeit“	5.41	0.96	34
Verständlichkeit „Buch selbst auswählen“	5.29	1.00	34
Verständlichkeit „Die Fragen sind voll schwer“	5.73	0.52	33
Verständlichkeit über alle 15 Vignetten	5.56	0.90	32

Anmerkung. Die beiden ersten Fragen zu Belastung und Akzeptanz beziehen sich auf alle fünfzehn Vignetten im Gesamten.

Die Studierenden geben ein mittleres Maß an Belastung durch das Ausfüllen des Vignettentests an. Die Akzeptanz sowie die Bewertungen der Verständlichkeit der Situationsbeschreibungen in den Vignetten werden durchgehend sehr hoch bewertet und können daher als zufriedenstellend angesehen werden.

Neben Akzeptanz, Belastung und Verständlichkeit sollten durch die Pilotstudie auch erste Ergebnisse bezüglich der Testeigenschaften gewonnen werden. So sollten die Verteilung des Testscores und die Itemschwierigkeiten untersucht werden. Selbstverständlich können diese Ergebnisse aufgrund des geringen *Ns* nur äußerst zurückhaltend interpretiert werden.

Um für die Studierendenantworten Testpunkte zu vergeben, werden Lösungsmuster für die jeweiligen Vignetten benötigt. Eine theoretische Ableitung der fachdidaktisch „richtigen“ Lösung für die jeweiligen Unterrichtssituationen erfolgte hier, nach dem Vorbild vorausgehender Studien, nicht (z. B. Tepner & Dollny, 2014). Für fachdidaktische Fragestellungen, die sich auf eine bestimmte Unterrichtssituation bzw. einen bestimmten Kontext beziehen, ist es nämlich sehr schwierig, eine eindeutig „richtige“ oder „falsche“ Lösung aus der Literatur abzuleiten (Schmelzing, Wüsten, Sandmann & Neuhaus, 2008). Die „richtige“ Lösung für die Vignetten wurde stattdessen durch die mehrstufige Expertenbefragung empirisch abgeleitet: Die durchschnittliche Antwort der Experten für eine Unterrichtssituation wird daher als die „richtige Lösung“ für eine Vignette festgesetzt (sog. „Expertenorm“ oder „aggregierter Experte“) (Oser & Forster-Heinzer, 2015; Oser, Heinzer & Salzmann, 2010; Seidel et al., 2010). Dieses Vorgehen schließt nicht nur an Vorarbeiten von renommierten Forschergruppen an, es berücksichtigt auch insbesondere, dass mit dem Vignettentest die impliziten Anteile des lesedidaktischen Wissens erfasst werden sollen. Implizite Wissensinhalte sind wie oben dargestellt nicht eindeutig kommunizierbar und liegen dem Handeln unbewusst zugrunde. Das bedeutet selbstredend, dass auch die Experten nicht, oder zumindest nicht eindeutig, kommunizieren können, weshalb eine bestimmte Handlungsalternative in einer gewissen Situation mehr oder weniger fachdidaktisch sinnvoll ist.

Zur Auswertung der Pilotstudie wurden die Ergebnisse der quantitativen Expertenbefragung als vorläufige Auswertungsnorm herangezogen. Dabei wurde mit dem Modalwert gearbeitet. Ein Item galt als gelöst (1 Punkt), wenn die Angabe

des Probanden mit dem Modalwert der Expertenbefragung übereinstimmte. Wenn der Proband eine Abweichung von $|x| = 1$ gekreuzt hatte, erhielt er für diese Antwort 0.5 Punkte. Wenn der Proband eine Abweichung von $|x| > 1$ gekreuzt hatte, erhielt er 0 Punkte. Die Probanden konnten zwischen 0 und 75 Punkten erreichen (da der Test 75 Items umfasste). Es zeigte sich, dass die Testscores der Probanden annähernd normalverteilt waren ($M = 28.6$, $SD = 2.7$). Allerdings war der Range der Testwerte als eingeschränkt zu bezeichnen (Range = 10, Min = 23, Max = 33). Ursächlich für den eingeschränkten Range könnte die kleine Stichprobe sein, in der nicht Studierende aller Semester repräsentiert waren und somit nicht alle möglichen Ausprägungen im Datensatz vorhanden sind.

Zur Analyse der Items wurden die Itemschwierigkeiten berechnet. Zu leichte (Itemschwierigkeit > 0.8) oder zu schwierige (Itemschwierigkeit < 0.2) wurden aus dem Test ausgeschlossen. Aufgrund der Ergebnisse der Pilotstudie wurden drei Vignetten aus dem Test ausgeschlossen, sodass das endgültige Testinstrument aus 12 Vignetten besteht.

Durch die Reduktion der Items und der Vignetten fiel ein Bereich der lesedidaktischen Forschung (aus dem heraus Vignetten entwickelt wurden) aus dem theoretischen Strukturmodell des lesedidaktischen Wissens heraus: der Bereich „Aufgabenschwierigkeit“. Es hat sich gezeigt, dass sich die Vignetten, die zu diesem Bereich entwickelt wurden, empirisch nicht bewährt haben.

Tabelle 5. Darstellung der Vignetten des finalen Testinstruments

	Modellfacette	
	Förderwissen/Intervention	Diagnose/Inhaltswissen
Lesestrategien	Gruppenarbeit zum Thema „Klimawandel“ (5)	Die Lesestrategien Mindmap (5)
	Die Fragen sind voll schwer! (5)	
Lesemotivation	„Ich kapiere die Bücher eh nie...“ (6)	
	Die Bücherkiste (5)	
Leseflüssigkeit	Lesetrainer und Lesesportler (4)	Wie kann die Leseflüssigkeit gesteigert werden? (5)
		Wählt euch ein Buch selbst aus! (5)
Textverstehen	„Ich verstehe den Satz nicht!“ (5)	
	Emil und die Detektive (5)	
Textschwierigkeit		Was machen Jugendliche in ihrer Freizeit? (5)
		Überprüfen des Textverständnisses (5)

Anmerkung. Die Anzahlen der Items pro Vignette können in den Klammern hinter dem Vignettentitel abgelesen werden.

9.3.4 Generierung der Auswertungsnorm

Durch die qualitative und quantitative Expertenbefragung sowie den Studienpilot wurden die zwölf geeignetsten Vignetten für das finale Testinstrument ausgewählt. Der abschließende Schritt der Testentwicklung bestand daher darin, eine finale Auswertungsnorm für das fertig gestellte Testinstrument zu erstellen. Das endgültige Testinstrument mit 12 Vignetten (60 Items) wurde dazu einer Gruppe von Experten der Lesedidaktik vorgelegt ($N = 15$). Es beteiligten sich insgesamt 10 Experten an der Erstellung der Auswertungsnorm (Alter: $M = 47.10$ Jahre, $SD = 7.53$, weiblich = 30 %, Professur: $N = 6$, Privatdozent: $N = 1$, Akademische Mitarbeiter: $N = 2$, keine Angabe: $N = 1$, Rücklauf: 66.7 %). Als Experten galten Personen, die sich hauptberuflich wissenschaftlich oder praktisch mit dem

Themenbereich Lesedidaktik beschäftigen. Des Weiteren war im Gegensatz zur vorausgehenden, umfangreichen Expertenbefragung ein wichtiges Kriterium bei der Expertenauswahl, dass diese im Bereich der Lesedidaktik besondere Qualifikationen auswiesen wie beispielsweise entsprechende Veröffentlichungen.

Die Experten kamen erwartungsgemäß nicht immer zu einem einstimmigen Urteil dahingehend, wie die Items der jeweiligen Unterrichtssituationen zu bewerten sind. Daher wurde die Einheitlichkeit, mit der die Experten die Items beurteilt haben, global untersucht. Hierzu wurde der Intraklassen-Korrelationskoeffizient (ICC) berechnet. Der Wert von $r = 0.87$ (95 %-Konfidenzintervall zwischen 0.82 und 0.92) spricht für eine gute Übereinstimmung zwischen den Experten (Wirtz & Caspar, 2002).

Um den aggregierten Experten und somit das richtige Lösungsmuster für eine Vignette zu erhalten, wurden die Rohwerte der Experten (Werte von 1 bis 6 auf der Likert-Skala pro Item) zunächst z -standardisiert, um Antworttendenzen zu vermeiden (siehe Tabelle 6 und Tabelle 7).

Tabelle 6. Expertenratings der Vignette „Überprüfen des Textverständnisses“

	Item 1	Item 2	Item 3	Item 4	Item 5
Experte 1	5	3	5	6	2
Experte 2	4	6	3	5	5
Experte 3	3	5	5	6	5
Experte 4	2	5	3	2	4
Experte 5	2	3	1	5	6
Experte 6	2	3	2	4	4
Experte 7	1	4	1	6	5
Experte 8	5	4	3	5	5
Experte 9	2	3	2	5	4
Experte 10	5	6	1	6	4

Anmerkung. Dargestellt sind die Rohwerte auf der sechsstufigen Likert-Skala. Siehe auch Beispielvignette in Abbildung 6.

Tabelle 7. Z-standardisierten Expertenratings der Vignette „Überprüfen des Textverständnisses“ und Median der Items

	Item 1	Item 2	Item 3	Item 4	Item 5
Experte 1	0.49	-0.73	0.49	1.10	-1.34
Experte 2	-0.53	1.23	-1.40	0.35	0.35
Experte 3	-1.64	0.18	0.18	1.10	0.18
Experte 4	-0.92	1.38	-0.15	-0.92	0.61
Experte 5	-0.68	-0.19	-1.16	0.77	1.25
Experte 6	-1.00	0.00	-1.00	1.00	1.00
Experte 7	-1.04	0.26	-1.04	1.13	0.69
Experte 8	0.67	-0.45	-1.57	0.67	0.67
Experte 9	-0.92	-0.15	-0.92	1.38	0.61
Experte 10	0.29	0.77	-1.64	0.77	-0.19
Median	-0.80	0.09	-1.02	0.89	0.61

Anmerkung. Der Median bezieht sich auf die z-standardisierten Itemratings über alle 10 Experten.

Im Anschluss wurde für die Antwortverteilung jedes Items der Median gebildet. Innerhalb einer Vignette wurden die Items dann mithilfe des Median in eine aufsteigende Reihenfolge gebracht. Das Item mit dem kleinsten Median ist laut durchschnittlicher Expertenmeinung die am wenigsten geeignete Handlungsalternative für die vorliegende Unterrichtssituation. Das Item mit dem höchsten Median ist laut durchschnittlicher Expertenmeinung die Handlungsalternative, die am geeignetsten ist.

Das Lösungsmuster, das sich hieraus für die Beispielvignette ergibt lautet: Item 3 < Item 1 < Item 2 < Item 5 < Item 4. Item 3 ist laut der durchschnittlichen Expertenmeinung, die am wenigsten geeignete Handlungsalternative für die vorgegebene Unterrichtssituation. Item 4 wurde von den Experten im Durchschnitt als die geeignetste Handlungsalternative beurteilt.

Die Punktvergabe für die Probanden erfolgt dann anhand eines Vergleichs der Antwortrelationen¹⁸ (Artelt, Beinicke, Schlagmüller & Schneider, 2009; Tepner & Dollny, 2014). Dazu werden die Antworten der Studierenden ebenfalls z-standardisiert und anhand des Median in eine aufsteigende Reihenfolge gebracht (siehe Tabelle 8).

Tabelle 8. Z-standardisierte Ratings zweier Probanden als Beispiel

	Item 1	Item 2	Item 3	Item 4	Item 5
Proband 1	-0.62	-0.10	-1.14	0.42	1.45
Proband 2	-0.77	0.19	-1.25	1.15	0.68

Bei diesem Vorgehen zur Punktvergabe werden immer zwei Items in Relation zueinander gesetzt (siehe Tabelle 9): Für dieses Beispiel wird zunächst überprüft, ob der aggregierte Experte Item 2 besser beurteilt als Item 1 (also ob $\text{Item 2} > \text{Item 1}$). Dann wird die Probandenantwort mit dieser Vorgabe abgeglichen. Wenn auch der Proband Item 2 besser als Item 1 bewertet hat, erhält er für diese Antwort einen Punkt, ansonsten 0 Punkte (siehe Tabelle 10) (modifiziert nach Tepner & Dollny, 2014).

Tabelle 9. Vergleich der Experten-Rangfolge der Items mit der Rangfolge der Probanden

	Item 3 mit Item 1	Item 3 mit Item 2	Item 3 mit Item 5	Item 3 mit Item 4	Item 1 mit Item 2	Item 1 mit Item 5	Item 1 mit Item 4	Item 2 mit Item 5	Item 2 mit Item 4	Item 5 mit Item 4
Experten	<	<	<	<	<	<	<	<	<	<
Proband 1	<	<	<	<	<	<	<	<	<	>
Proband 2	<	<	<	<	<	<	<	<	<	<

¹⁸ Dieses Vorgehen in der Auswertung unterscheidet sich von der Auswertung in der Pilotstudie: In der Pilotstudie war eine Auswertung anhand von Antwortrelationen nicht möglich, da zu diesem Zeitpunkt die endgültige Anzahl an Items pro Vignette noch nicht feststand.

Tabelle 10. Punktvergabe für die Studierenden anhand des Vergleichs der Antwortrelationen der Experten und der Probanden

	Item 3 mit Item 1	Item 3 mit Item 2	Item 3 mit Item 5	Item 3 mit Item 4	Item 1 mit Item 2	Item 1 mit Item 5	Item 1 mit Item 4	Item 2 mit Item 5	Item 2 mit Item 4	Item 5 mit Item 4	Summe
Proband 1	1	1	1	1	1	1	1	1	1	0	9
Proband 2	1	1	1	1	1	1	1	1	1	1	10

Es spielt bei diesem Vorgehen der Punktevergabe also keine Rolle, auf welcher exakten Position der Likert-Skala die Probanden die Items bewertet haben; für die Bewertung der Studierendenantworten ist ausschließlich entscheidend, ob die Relation zwischen zwei Items korrekt gebildet wurde. Es wird hier also die Fähigkeit der Studierenden bewertet, verschiedene Handlungsmöglichkeiten gegeneinander abwägen zu können. Eben diese Fähigkeit ist auch für das Unterrichten relevant, denn hier sind Lehrkräfte häufig damit konfrontiert, zwischen verschiedenen konkurrierenden Handlungsalternativen wählen zu müssen.

Bei einer Vignette mit 5 Items können insgesamt 10 Relationen gebildet werden, d. h. der Proband kann innerhalb dieser Vignette 0 bis 10 korrekte Relationen erreichen. Je mehr korrekte Relationen ein Proband innerhalb einer Vignette erreicht, desto ähnlicher ist seine Antwort der „optimalen“ Antwort des aggregierten Experten. Eine höhere Punktzahl spricht daher für eine höhere Ausprägung auf dem Zielkonstrukt, dem lesedidaktischen Wissen.

Die korrekt gebildeten Itemrelationen werden also pro Vignette addiert, sodass für jede Vignette ein „Vignettensummenscore“ resultiert. Dieses Vorgehen wird in der Literatur als *score-basierter Ansatz* bezeichnet (Cook, Dodd & Fitzpatrick, 1999) und bietet sich für das vorliegende Testinstrument an, da dieses eine sog. Testletstruktur aufweist (Bühner, 2011; Wainer & Kiely, 1987). Eine Testletstruktur liegt dann vor, wenn Items innerhalb eines Testinstruments gruppiert vorliegen. Im vorliegenden Fall beziehen sich immer mehrere Items auf eine gemeinsame Vignette und damit einhergehend auf eine gemeinsame Fragestellung, wodurch eine lokale stochastische Abhängigkeit zwischen den Items resultieren kann. Eine Nichtbeachtung der Testletstruktur von Testinstrumenten kann sich verfälschend auf die Schätzung von psychometrischen Gütekriterien wie Reliabi-

liät, Trennschärfe oder Itemschwierigkeit auswirken (Eckes, 2015). Des Weiteren liegt für die gebildeten Itemrelationen innerhalb einer Vignette eine inhaltliche Abhängigkeit vor, da immer zwei Items in Bezug zueinander gesetzt werden. Die Verwendung des score-basierten Ansatzes erscheint für die Auswertung des hier vorgestellten vignettenbasierten Testverfahrens daher als besonders angemessen.

10 Methode

Im folgenden Kapitel wird der Verlauf der Hauptstudie sowie die geplante Auswertung der Daten beschrieben. Die Haupterhebung fand im Mai und Juni 2015 an den Pädagogischen Hochschulen Heidelberg, Ludwigsburg, Karlsruhe und Schwäbisch Gmünd, der Universität Heidelberg und den Staatlichen Seminaren für Lehrerbildung Ludwigsburg und Heppenheim statt.

10.1 Stichprobe

Es wurden insgesamt 581 Deutschlehramtsstudierende mit dem Testinstrument befragt (Alter: $M = 23.08$ Jahre, $SD = 3.32$, weiblich: 87.8 %). Die Stichprobe setzte sich aus Studierenden verschiedener Deutschlehramtsstudiengänge zusammen: Primarstufe $N = 313$, Sekundarstufe $N = 144$, Sonderpädagogik $N = 77$, Gymnasiallehramt $N = 47$ ¹⁹.

Zusätzlich wurden zur Untersuchung der Validität des Testverfahrens drei Kontrollgruppen erhoben: Germanistikstudierende ($N = 33$, Alter: $M = 21.72$ Jahre, $SD = 2.22$, weiblich: 78.8 %), Psychologiestudierende ($N = 67$, Alter: $M = 21.33$ Jahre, $SD = 2.39$, weiblich: 86.6 %) und Lehramtsstudierende fachfremder Lehrämter (ohne das Fach Deutsch) ($N = 69$, Alter: $M = 22.43$ Jahre, $SD = 2.83$, weiblich: 59.4 %).

¹⁹ An den Pädagogischen Hochschulen (an denen hauptsächlich erhoben wurde) werden Seminare zur Lese- bzw. Deutschdidaktik studiengangübergreifend angeboten; daher sind in der Stichprobe Deutschlehramtsstudierende verschiedener Lehrämter vertreten.

Tabelle 11. Beschreibung der verschiedenen Substichproben

Substichprobe	N	Alter		% weiblich	Studienabschnitt ^a								Ref
		M	SD		1	2	3	4	5	6	7	8+	
Deutschstudierende Primarstufe	311	22.85	2.81	92.6	73	67	9	22	5	74	26	35	--
Deutschstudierende Sekundarstufe	143	23.96	3.86	82.5	20	31	10	22	8	17	4	10	20
Deutschstudierende Sonderpädagogik	77	22.85	4.32	87.0	21	35	2	9	--	4	1	--	5
Deutschstudierende Gymnasiallehramt	47	22.39	2.41	72.3	5	13	2	10	2	6	1	8	--
Studierende fachfremder Lehrämter	69	22.43	2.83	59.4	24	30	7	2	--	3	--	3	--
Psychologiestudierende	67	21.33	2.39	86.6	--	53	--	9	--	2	--	1	--
Germanistikstudierende	33	21.72	2.22	78.8	1	12	1	10	1	3	--	--	--

Anmerkungen. ^aDie Variable Studienabschnitt bezieht sich auf die Anzahl an Fachsemestern.

„8+“ beschreibt Studierende, die im 8. oder einem höheren Semester studieren, „Ref“ bezeichnet das Referendariat.

Es handelte sich um eine anfallende Stichprobe (auch: Gelegenheitsstichprobe). Das bedeutet, dass die Studierenden im Kontext von Lehrveranstaltungen zufällig befragt wurden. Dies führt zu den teilweise sehr unterschiedlich stark besetzten Zellen.

10.2 Verlauf der Hauptstudie

Die Studierendenbefragungen wurden von zwei geschulten Testleiterinnen durchgeführt. Der Ablauf der Befragungen war standardisiert: Zunächst wurde mit Dozenten aus dem Fachbereich Deutsch der jeweiligen Hochschule Kontakt aufgenommen und angefragt, ob sie bereit wären, die Befragung im Rahmen einer Seminarsitzung durchführen zu lassen. Wenn die Dozenten zustimmten, fand die Befragung an einem von ihnen vorgeschlagenen Termin statt. Den Studierenden wurde die Befragung in der Regel nicht angekündigt. Es stand den Studierenden selbstverständlich frei, nicht an der Befragung teilzunehmen.

Mit einer kurzen PowerPoint-Präsentation wurden die Studierenden in die Bearbeitung des Fragebogens eingeführt. Die Studierenden wurden darauf hingewiesen, dass die Teilnahme an der Befragung anonym ist und dass eine Teilnahme bzw. die Nicht-Teilnahme an der Befragung keine Auswirkungen auf ihre Leistungsbeurteilung hat. Zur Bearbeitung der Vignetten wurde den Studierenden mitgeteilt, dass sie sich zunächst alle Items einer Vignette durchlesen und erst danach ihre Antwort abgeben sollen. Diese Information stand zusätzlich im Testheft.

Als Belohnung für die Teilnahme an der Befragung erhielten die Studierenden eine Süßigkeit. Es gab ferner die Möglichkeit an einer Verlosung von Gutscheinen teilzunehmen. Die Voraussetzung hierzu war, dass die Studierenden bereit waren, sich in einer separaten Liste mit ihrem Namen und ihrer (privaten) E-Mail-Adresse einzutragen, da eine Verlosung ansonsten nicht möglich gewesen wäre.

Nachdem die Studierendenbefragungen abgeschlossen waren, wurden die Daten aller vorliegenden Fragebögen durch sechs geschulte Mitarbeiter in eigens dafür erstellte Excel-Masken eingegeben ($N = 769$ Fragebögen). Da das Testheft in acht unterschiedlich rotierten Versionen (verwendete Instrumente siehe 10.3) eingesetzt wurde, gab es zu jeder Version ein Kodiermanual, das den Mitarbeitern ausgehändigt wurde.

Nach der Eingabe der Fragebögen wurden 135 zufällig ausgewählte Fragebögen ein zweites Mal eingegeben (17.6 %), um auf mögliche Fehler bei der ersten Eingabe aufmerksam zu werden. Die eingegebenen Daten der Erst- und Zweiteingabe wurden miteinander korreliert, um deren Übereinstimmung zu überprüfen. Das Ergebnis lässt auf eine sehr zuverlässige Eingabe der Daten schließen ($r = 0.99$, $p \leq 0.001$).

Im Anschluss daran wurden die Daten aus den verschiedenen Fragebogenversionen in einen SPSS-Datensatz überführt. Dies geschah durch zwei geschulte Mitarbeiterinnen. Abschließend wurde der finale SPSS-Datensatz nochmals gründlich kontrolliert und mit einzelnen, zufällig ausgewählten Fragebögen abgeglichen.

10.3 Verwendete Instrumente

Für die Haupterhebung wurde ein Testheft erstellt. Dieses lag in acht unterschiedlich rotierten Versionen vor, um Ermüdungs- und/oder Reihenfolgeeffekte zu

vermeiden. Außerdem sollte so verhindert werden, dass die Studierenden voneinander abschreiben.

Zu Beginn der Befragung wurden die Studierenden gebeten, einen individuellen persönlichen Code zu generieren (erste beide Buchstaben des Vornamens der Mutter, GeburtsTAG, erste beide Buchstaben des Vornamens des Vaters, Geschlecht). Dieser Code sollte dazu dienen, die Datensätze der Studierenden verbinden zu können, falls diese mehrfach befragt werden würden (im weiteren Verlauf der Studie angedacht). Neben den Vignetten wurden im Testheft weitere soziodemographische Angaben und Konstrukte abgefragt, die im Folgenden dargestellt werden.

10.3.1 Soziodemographische Angaben

Die Studierenden wurden gebeten, einige soziodemographische Angaben zu machen: Es wurden Alter, Geschlecht, Studienort, Semesteranzahl, Studiengang, Fächerkombination, in welchem Bundesland bzw. Land das Abitur abgelegt wurde, Bildung der Eltern, relevante Nebentätigkeiten (z. B. Nachhilfe geben) und vorherige Studiengänge/Berufsausbildungen abgefragt.

Daneben wurden die Studierenden gebeten anzugeben, ob sie bereits Praxiserfahrung im Deutsch- bzw. im Leseunterricht gesammelt haben (im Integrierten Semesterpraktikum, als Aushilfslehrer, Sonstiges). Es wurde außerdem gefragt, ob die Studierenden bereits Seminare zur Lesedidaktik besucht haben und wenn ja.

10.3.2 Vignettentest zum schreibdidaktischen Wissen

Neben dem Vignettentest zur Erfassung des lesedidaktischen Wissens wurde in der Befragung auch ein Vignettentest zur Erfassung des schreibdidaktischen Wissens eingesetzt, der in einem „Schwesterprojekt“ im FuN-Kolleg EKoL entwickelt wurde. Dieser Vignettentest besteht ebenfalls aus zwölf Vignetten mit jeweils fünf Items, die ebenfalls auf einer sechsstufigen Likert-Skala zu beantworten waren.

In weiteren Analysen soll es so möglich sein, Zusammenhänge zwischen dem lese- und dem schreibdidaktischen Wissen untersuchen zu können. Diese Analysen sind jedoch nicht Gegenstand der vorliegenden Arbeit.

10.3.3 Abiturdurchschnitt und Sprachliches Denken

Mit dem vorliegenden Testinstrument soll lesedidaktisches Wissen, ein professioneller Wissensinhalt von Lehrpersonen bzw. Lehramtsstudierenden, erfasst werden. Bei der Überprüfung der Validität ist wichtig zu untersuchen, ob das lesedidaktische Wissen durch das Testinstrument von anderen, verwandten Konstrukten abgegrenzt werden kann (Überprüfung der diskriminanten Validität). Da die allgemeine Intelligenz als ein wichtiger Prädiktor für verschiedene Arten von kognitiven Leistungen gilt (Hülshager & Maier, 2008; Neisser et al., 1996), ist es für die psychometrische Beurteilung des Testinstruments wichtig, dieses u. a. von allgemeiner Intelligenz abzugrenzen. So liegen empirische Befunde vor, dass die allgemeine kognitive Leistung (oftmals operationalisiert über die Abiturnote) bedeutsam für den Wissensaufbau von Lehramtsstudierenden ist (König, Tachtoglou & Seifert, 2012). Ferner wurde gezeigt, dass die Abiturnote von Lehramtsstudierenden mit deren fachbezogenem Wissen zusammenhängt (Blömeke, Felbrich & Müller, 2008). Es ist daher sinnvoll, die Abiturnote als Marker für allgemeine kognitive Leistung in entsprechenden Analysen zu kontrollieren (Kunina-Habenicht et al., 2013).

Neben einem Maß für die allgemeine kognitive Leistungsfähigkeit (operationalisiert über die Abiturnote) wurde eine Skala zum sprachlichen Denken²⁰ eingesetzt (Kersting, Althoff & Jäger, 2008), da das Testinstrument durch die Verwendung von Text-Vignetten sehr leseintensiv ist.

Die Skala zum sprachlichen Denken besteht aus 20 Wortanalogien. Es ist jeweils eine Wortanalogie vorgegeben und eine zugehörige zweite Wortanalogie soll nach deren Vorbild vervollständigt werden (Beispiel: Schaf : Wolle = Vogel : ? a) Flügel, b) Nest, c) Adler, d) Federn, E) Tier). Die Studierenden konnten jeweils aus fünf möglichen Antworten auswählen. Für die Bearbeitung der 20 Wortanalogien hatten die Studierenden laut Vorgabe aus dem Testmanual 4:30 Minuten Zeit. Die Wortanalogien wurden im Testheft daher in jeder Version des Testhefts an erster

²⁰ Die eingesetzte Skala wird von den Autoren als „Sprachliches Denken und schlussfolgerndes Denken“ bezeichnet (Kerstin, Althoff & Jäger, 2008). Zur vereinfachten Lesbarkeit wird diese Skala hier „Sprachliches Denken“ genannt.

Stelle präsentiert, da nur so die Zeitmessung durch die Testleiterin erfolgen konnte. Es ergab sich ein Cronbachs α von 0.79²¹.

10.3.4 PK

Neben PCK sind auch die Wissensdomänen PK und CK in der Forschung zum professionellen Wissen von Lehrkräften von Interesse. Um neben PCK einen weiteren professionellen Wissensbereich abzufragen, wurde ein Test zur Erfassung von PK eingesetzt (König & Blömeke, 2010). Das Testinstrument umfasst offene sowie geschlossene Fragen. In der vorliegenden Studie wurden aus zeitökonomischen Gründen ausschließlich die geschlossenen Fragen des Testinstruments eingesetzt. Dies waren insgesamt 10 Fragen im Multiple-Choice-Format aus den Bereichen: Umgang mit Heterogenität (1 Frage), Strukturierung (1 Frage), Klassenführung (3 Fragen), Motivierung (2 Fragen) und Leistungsbeurteilung (3 Fragen). Innerhalb der 10 Fragen zu PK konnten die Probanden maximal 30 Punkte erreichen, da es innerhalb der Fragen z.T. mehrere Teilaufgaben gab, die einzeln bewertet wurden. Es ergab sich ein Cronbachs α von 0.81.

10.3.5 Lesedidaktische Überzeugungen

Da für die Unterrichtsqualität auch die Überzeugungen von Lehrkräften bedeutsam sind, wurden für den Leseunterricht spezifische Überzeugungen abgefragt (Behrmann & Souvignier, 2013). Die Skalen beziehen sich auf Überzeugungen zur Instruktion bei Leseaufgaben. Es wurden vier Subskalen abgefragt: Konstruktivistische Überzeugung (z. B. „Um einen kompetenten Umgang mit Texten zu erlernen hilft es, Schülerinnen und Schüler ihre eigenen Herangehensweisen diskutieren zu lassen.“, 6 Items), direkt-transmissive Überzeugung (z. B. „Man sollte von Schülerinnen und Schülern verlangen, beim Bearbeiten von Texten so vorzugehen, wie es im Unterricht erarbeitet wurde.“, 6 Items) und Globale/Allgemeine Überzeugung (z. B. „Die Unterrichtsinhalte sollten den Schülerinnen und Schülern vorrangig praktische Lernerfahrungen ermöglichen.“, 6 Items). Die Items waren jeweils auf einer fünfstufigen Likert-Skala zu beantworten (1 = *Trifft gar nicht zu*, 5 = *Trifft völlig zu*).

²¹ Für jede der eingesetzten Skalen wird als Marker für die Reliabilität der Skala das interne Konsistenzmaß Cronbachs α angegeben (Bühner, 2011).

Außerdem wurde die Skala zur Verwendung expliziter Lesestrategien im Unterricht eingesetzt (Behrmann & Souvignier, 2013). Hier sollten die Studierenden auf einer siebenstufigen Skala zwischen zwei gegensätzlichen Positionen zur Strategievermittlung ankreuzen (z. B. „Für das Textverständnis ist es wichtig, Schülerinnen und Schülern zu erklären, warum es sinnvoll ist, beim Lesen eines Textes zunächst wichtige Informationen zu unterstreichen.“ versus „Schülerinnen und Schüler finden auch von alleine heraus, dass sie beim Lesen eines Textes zunächst die wichtigsten Informationen unterstreichen sollten, um ihr Textverständnis zu verbessern.“, 7 Items). Es ergaben sich folgende Werte für Cronbachs α : Konstruktivistische Überzeugung $\alpha = 0.73$, direkt-transmissive Überzeugung $\alpha = 0.74$, globale/allgemeine Überzeugung $\alpha = 0.73$, Skala zur Verwendung expliziter Lesestrategien $\alpha = 0.81$.

10.3.6 Berufsbezogenes Selbstkonzept

Es wurden drei Skalen zum berufsbezogenen Selbstkonzept angehender Lehrkräfte abgefragt (Retelsdorf et al., 2014a): Selbstkonzept Fach²² (z. B. „Die Inhalte meines Studienfachs Deutsch bereiten mir keine Schwierigkeiten.“, 3 Items), Selbstkonzept Erziehung (z. B. „Ich kann gut mit Kindern und Jugendlichen umgehen.“, 4 Items) und Selbstkonzept Diagnostik (z. B. „Ich kann Leistungen anderer gut beurteilen.“, 4 Items). Die Items waren jeweils auf einer fünfstufigen Likert-Skala zu beantworten (1 = *Trifft gar nicht zu*, 5 = *Trifft völlig zu*). Für die Analysen wurde für jede Subskala der Mittelwert aus den zugehörigen Items gebildet. Es ergaben sich folgende Werte für Cronbachs α : Selbstkonzept Fach $\alpha = 0.79$, Selbstkonzept Erziehung $\alpha = 0.74$, Selbstkonzept Diagnostik $\alpha = 0.74$.

10.3.7 Berufswahlmotive

Es wurde ein Fragebogen zu den Berufswahlmotiven der Probanden eingesetzt (Pohlmann & Möller, 2010). Dieser Fragebogen umfasst sechs Subskalen: Nützlichkeit (z. B. „Ich habe das Lehramtsstudium gewählt, weil ich als Lehrer/in gut verdiene.“, 8 Items), Pädagogisches Interesse (z. B. „Ich habe das Lehramtsstudium gewählt, weil ich gerne mit Kindern und Jugendlichen arbeite.“, 6 Items), Fä-

²² Die Skala „Selbstkonzept Fach“ wurde auf das Fach Deutsch angepasst, das sich das Selbstkonzept der Studierenden bezüglich der verschiedenen Fächer, die sie studieren, unterscheiden könnte.

higkeitsüberzeugung (z. B. „Ich habe das Lehramtsstudium gewählt, weil ich gut erklären kann.“, 5 Items), Soziale Einflüsse (z. B. „Ich habe das Lehramtsstudium gewählt, weil mir von Freunden und Bekannten zum Lehramtsstudium geraten wurde.“, 5 Items), Geringe Schwierigkeit (z. B. „Ich habe das Lehramtsstudium gewählt, weil ich denke, dass ich in diesem Studium ohne große Mühe durchkomme.“, 4 Items) und Fachliches Interesse (z. B. „Ich habe das Lehramtsstudium gewählt, weil ich mich gern mit den Inhalten meiner Fächer beschäftige.“, 5 Items). Die Items waren jeweils auf einer fünfstufigen Likert-Skala zu beantworten (1 = *Trifft gar nicht zu*, 5 = *Trifft völlig zu*). Für die Analysen wurde für jede Subskala der Mittelwert aus den zugehörigen Items gebildet. Es ergaben sich folgende Werte für Cronbachs α : Nützlichkeit $\alpha = 0.90$, Pädagogisches Interesse $\alpha = 0.83$, Fähigkeitsüberzeugung $\alpha = 0.78$, Soziale Einflüsse $\alpha = 0.83$, Geringe Schwierigkeit $\alpha = 0.87$, Fachliches Interesse $\alpha = 0.94$.

10.3.8 BIG-5

Zur Erfassung der Persönlichkeitseigenschaften der Probanden wurden Extraversion („Ich bin jemand, der kommunikativ, gesprächig ist.“), Neurotizismus („Ich bin jemand, der sich oft Sorgen macht.“), Gewissenhaftigkeit („Ich bin jemand, der gründlich arbeitet.“), Verträglichkeit („Ich bin jemand, der verzeihen kann.“) und Offenheit („Ich bin jemand, der originell ist, neue Ideen einbringt.“) abgefragt. Dazu wurde eine Kurzskala der BIG-5 verwendet (Gerlitz & Schupp, 2005). Die Items waren auf einer fünfstufigen Likert-Skala zu beantworten (1 = *Trifft gar nicht zu*, 5 = *Trifft völlig zu*). Pro Subskala wurden 3 Items abgefragt. Es ergaben sich folgende Werte für Cronbachs α : Extraversion $\alpha = 0.65$, Neurotizismus $\alpha = 0.50$, Gewissenhaftigkeit $\alpha = 0.67$, Verträglichkeit $\alpha = 0.55$, Offenheit $\alpha = 0.65$.

10.4 Analysen

In einem ersten Analyseschritt²³ soll eine konfirmatorische Faktorenanalyse berechnet werden, um die theoretisch angenommene zweidimensionale Struktur der Daten zu überprüfen (Muthén & Muthén, 2012). *Mplus* stellt dabei mehrere Fit-

²³ Es wurde jeweils die gängige Software für den jeweiligen Analysetyp gewählt: Mittelwertvergleiche, deskriptive Analysen, etc.: SPSS, Analysen zur Datenstruktur: *Mplus*, Skalierung: verschiedene R-Packages.

Indices zur Verfügung, die eine Beurteilung der allgemeinen Modellgüte ermöglichen. Ein χ^2/df -Verhältnis zwischen 1 und 3 (Jöreskog & Sörbom, 1993), ein CFI²⁴-Wert über 0.9 und ein RMSEA²⁵-Wert von unter 0.6 weisen auf einen guten Modellfit hin (Hu & Bentler, 1998).

Um differenzierte psychometrische Aussagen über die einzelnen Testaufgaben treffen zu können, soll der Vignettentest im Anschluss an die Überprüfung der Datenstruktur mit den R-Packages *TAM* (Kiefer, Robitzsch & Wu, 2016) und *eRm* (Mair, Hatzinger, Maier & Rusch, 2015) mit dem Partial Credit Modell skaliert werden (Masters, 1982). Das Partial Credit Modell ist eine Verallgemeinerung des Rasch-Modells für polytome Items mit geordneten Antwortkategorien. Hier werden Schwellenparameter definiert (Anzahl Schwellenparameter = Anzahl Antwortkategorien – 1), die den Übergang von einer Kategorie in die nächsthöhere darstellen. Die Schwellenparameter sollten entlang des latenten Merkmals geordnet auftreten, da mit zunehmender Ausprägung des latenten Konstruktes die Wahrscheinlichkeit steigen sollte, eine höhere Kategorie, d. h. eine höhere Punktzahl, zu erreichen.

Die Passung der einzelnen Vignetten zum Modell sowie die globale Passung sollen mit der Infit- und Outfit-Statistik untersucht werden (Überblick: Ames & Penfield, 2015; Wright & Panchapakesan, 1969). Werte nahe 1 sprechen für eine gute Passung der Items zum Modell; bei Werten von über 1.5 bzw. unter 0.5 sollte über einen Ausschluss der Items aus dem Modell nachgedacht werden, da die Modellpassung in solchen Fällen zweifelhaft ist (Ayala, 2009). Andere Autoren nehmen hier strengere Kriterien an und sprechen bei Fit-Indices von 0.8 bis 1.2 von einer guten Modellpassung (Adams & Wu, 2002).

Lokale stochastische Abhängigkeiten zwischen den Vignetten bzw. den Items innerhalb der Vignetten sollen mit der Q3-Statistik untersucht werden (Yen, 1984; Yen, 1993). Werte von über 0.2 indizieren lokale Abhängigkeit. Zwischen den Itemrelationen der einzelnen Vignetten wird eine lokale Abhängigkeit erwartet (da Testletstruktur, siehe 9.3.4); zwischen den Vignetten(-summenscores) wird keine lokale Abhängigkeit erwartet

²⁴ „Comparative Fit Index“

²⁵ „Root Mean Square Error of Approximation“

Des Weiteren sollen Graphiken erstellt werden, um die Positionen der Itemkategorieschwellen und der Personenfähigkeiten in derselben Abbildung zu illustrieren. Diese Graphiken werden in der Literatur als Wright Maps (Irribarra & Freund, 2015) oder Person-Item Maps (Mair et al., 2015) bezeichnet.

Außerdem soll die EAP-Reliabilität des Vignettentests berechnet werden (Rost, 2004). Dieses Reliabilitätsmaß kann Werte zwischen 0 und 1 annehmen, wobei höhere Werte für ein höheres Ausmaß an Reliabilität sprechen. Die Werte der EAP-Reliabilität sind mit Werten für Cronbachs α vergleichbar.

Zur Überprüfung der Validität des Vignettentests soll mit der Statistiksoftware SPSS ein Kontrollgruppenvergleich mit einer einfaktoriellen ANCOVA durchgeführt werden (IBM Corporation, 2013). Als Kovariaten sollen hier die beiden Indikatoren für allgemeine kognitive Leistungsfähigkeit (Abiturdurchschnitt, Sprachliches Denken) aufgenommen werden.

Des Weiteren soll in SPSS mit einer einfaktoriellen ANOVA untersucht werden, ob ein (querschnittlicher) Anstieg des lesedidaktischen Wissens der Deutsch-Studierenden über die Semester beobachtet werden kann.

In latenten Regressionsmodellen soll mit *Mplus* untersucht werden, ob die Studiendauer, die Lehrerfahrung und der Besuch lesedidaktischer Seminare das lesedidaktische Wissen der Studierenden vorhersagen kann (Muthén & Muthén, 2012). Weiterhin soll untersucht werden, inwiefern die Studiendauer und die praktischen Erfahrungen der Studierenden inkrementelle Prädiktoren des lesedidaktischen Wissens darstellen. In diesen latenten Regressionsmodellen gelten – ebenso wie bei der konfirmatorischen Faktorenanalyse – ein χ^2/df -Verhältnis zwischen 1 und 3, ein CFI-Wert über 0.9 und ein RMSEA-Wert von unter 0.6 als Indikatoren für einen guten Modellfit.

Der Zusammenhang zwischen dem lesedidaktischen Wissen und den verwandten Konstrukten soll durch bivariate Korrelationen in SPSS überprüft werden.

10.4.1 Umgang mit fehlenden Werten

Zunächst wurde der Datensatz auf systematischen Dropout untersucht. Es konnte festgestellt werden, dass bei keiner Skala ein systematischer Dropout vorliegt, so dass von MCAR (Missing Completely at Random) ausgegangen werden kann

(MCAR-Test nach Little: Sprachliches Denken: $\chi^2(19) = 0.823$, $p > 0.05$, Vignettentest: $\chi^2(3449) = 3525.866$, $p > 0.05$, PK: $\chi^2(2599) = 2638.251$, $p > 0.05$, alle anderen Skalen ebenfalls $p > 0.05$) (Little, 1988).

Probanden, die über 50 % fehlende Werte innerhalb des Vignettentests aufwiesen, wurden aus den weiteren Analysen ausgeschlossen. In diesen Fällen konnte nicht von einer ernsthaften Bearbeitung der Unterrichtsvignetten ausgegangen werden, was eine zentrale Voraussetzung für die Analysen zur Datenstruktur und die Skalierung darstellt. Fünf Probanden mussten aufgrund dieses Kriteriums aus den Analysen ausgeschlossen werden, sodass sich die nachfolgenden Berechnungen auf den Stichprobenumfang $N = 764$ beziehen.

Fehlende Werte wurden durch Imputation bzw. durch Vergabe von 0 Punkten ersetzt. Dabei wurde folgendermaßen vorgegangen:

Fehlende Werte bei der Skala zum sprachlichen Denken und zu PK wurden mit 0 Punkten bewertet: Bei der Skala zum sprachlichen Denken handelt es sich um einen Speedtest; die Skala zu PK wurde an letzter Stelle im Testheft abgefragt, sodass die Probanden bei der Bearbeitung dieser Skala ebenfalls unter Zeitdruck standen. Es kann daher davon ausgegangen werden, dass fehlende Werte aufgrund mangelnder Bearbeitungszeit auftreten. Dies kann als Indikator für fehlende Kompetenz angesehen und daher mit 0 Punkten bewertet werden (Bühner, 2011). Für beide Skalen resultierten daher vollständige Variablen.

Der Vignettentest, die Skalen zur berufsbezogenen Selbstwirksamkeit, zu den Berufswahlmotiven, zu den BIG-5 und zur Einstellung im Leseunterricht wurden mit dem EM-Verfahren in SPSS unter Rückgriff auf weitere relevante Variablen (Beanspruchung und Akzeptanz bezogen auf die Vignetten, Geschlecht) imputiert (Allison, 2002; IBM Corporation, 2013).

10.4.2 Überprüfung der Voraussetzungen der Analysen

Neben einem ausreichend großen $N > 300$ (Marsh, Hau, Balla & Grayson, 1998; Myers, Ahn & Jin, 2011) ist das Vorliegen einer multivariaten Normalverteilung in den Daten eine wichtige Voraussetzung für die Berechnung von konfirmatorischen Faktorenanalysen (Backhaus, Erichson & Weiber, 2015). Da die Analyse der Datenstruktur des Vignettentests ein zentrales Forschungsanliegen darstellt,

wurde die multivariate Normalverteilung anhand des Mardia's Test untersucht (Mardia, 1970, 1974). Mardia's Test wurde mithilfe des R-Packages *MVN* berechnet (Korkmaz, Goksuluk & Zararsiz, 2015). Die Testwerte zur multivariaten Wölbung und zur multivariaten Kurtosis weisen auf eine Verletzung der multivariaten Normalverteilung hin (alle $ps \leq 0.05$). In Abbildung 9 wird deutlich, dass es einige Ausreißer gibt, die zur Verletzung der multivariaten Normalverteilung beitragen.

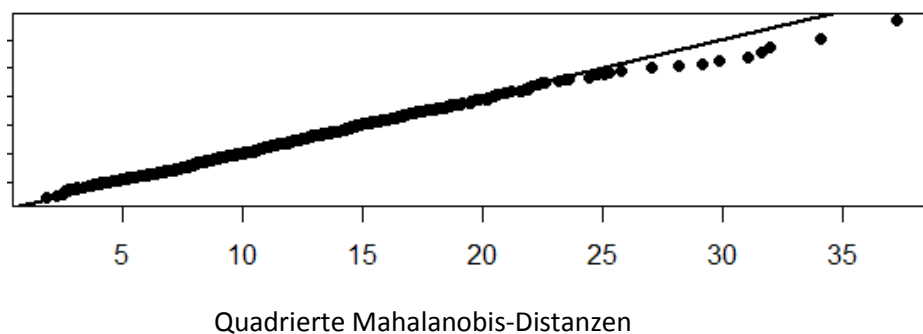


Abbildung 9. Darstellung der quadrierten Mahalanobis-Distanzen.

Da für den Vignettentest nicht von einer multivariaten Normalverteilung der Testwerte ausgegangen werden kann, wird bei der Berechnung der konfirmatorischen Faktorenanalyse ein Schätzer („ULSMV“; Muthén & Muthén, 2012, S. 603) verwendet, der ohne die Annahme der multivariaten Normalverteilung auskommt (Backhaus et al., 2015, S. 169).

10.4.3 Hierarchische Datenstruktur

In der Hauptuntersuchung wurden Deutschlehramtsstudierende aus verschiedenen Hochschulen befragt. Daher sollte überprüft werden, ob die Zugehörigkeit zu einer gewissen Hochschule die Testleistung beeinflusst. Dies wurde nicht erwartet, da die Vignetten unabhängig von einem spezifischen Hochschulcurriculum erstellt wurden. Mit *Mplus* wurde der Intraklassen-Korrelationskoeffizient (ICC) berechnet (Muthén & Muthén, 2012). Dieser Kennwert ermöglicht die Schätzung, wie viel der Varianz der abhängigen Variable, im vorliegenden Fall des lesedidaktischen Wissens, durch die Zugehörigkeit zu einer bestimmten Gruppe, im vorlie-

genden Fall einer gewissen Hochschule, bestimmt wird. Der ICC von 0.011 ($p \leq 0.05$) spricht dafür, dass die Zugehörigkeit zu einer bestimmten Hochschule tatsächlich keinen Einfluss auf die Testleistung hat.

11 Ergebnisse

Im Anschluss an die Darstellung der deskriptiven Analysen werden die Ergebnisse zur Überprüfung der Datenstruktur, zur Skalierung, zum Kontrollgruppenvergleich, zum (querschnittlichen) Anstieg des lesedidaktischen Wissens über die Studiensemester sowie zum Einfluss der Lehrerfahrung/lesedidaktischer Seminare auf das Abschneiden im Vignettentest präsentiert. Abschließend werden korrelative Zusammenhänge des lesedidaktischen Wissens zu PK, lesedidaktischen Überzeugungen, Berufswahlmotivation, Aspekten des berufsbezogenen Selbstkonzeptes sowie zu den BIG-5 berichtet.

11.1 Deskriptive Analysen²⁶

Die Vignetten umfassen in der Regel 5 Items²⁷. Pro Vignette konnten daher zehn Itemrelationen gebildet werden, für die Testpunkte vergeben wurden. Insgesamt sind pro Vignette somit 0 bis 10 Punkte erreichbar. Eine 11-stufige Gliederung des zugrunde liegenden latenten Konstrukts „lesedidaktisches Wissen“ ist allerdings wenig sinnvoll: Zum einen ist es äußerst schwierig, zwischen solch geringen Abstufungen statistisch wie inhaltlich differenzieren zu können. Zum anderen kann erwartet werden, dass die Stichprobe hinsichtlich der zu erfassenden Fähigkeit relativ homogen ist (da z. B. keine erfahrenen Lehrkräfte befragt werden). Daher sollten breitere Punktkategorien gebildet werden, die mehrere Relationen zusammenfassen.

Die Kategorienbildung sollte sowohl auf theoretisch-logischen Überlegungen als auch auf empirischen Beobachtungen basieren. Zunächst ist Folgendes zu überlegen: Wenn ein Proband die am besten geeignete Handlungsalternative innerhalb einer Vignette erkennt und das Kreuz daher am weitesten rechts auf der Likert-Skala setzt (Anker am Skalenende auf der rechten Seite: 6 = *Trifft voll und ganz zu*), dann sind automatisch vier Relationen richtig gekreuzt, unabhängig davon, in

²⁶ Die deskriptiven Analysen sowie die Analysen zur Datenstruktur, zu Skalierung, zur Praxiserfahrung und zu Zusammenhängen zu den Kovariaten beziehen sich ausschließlich auf die Hauptstichprobe, d. h. die Deutschlehramtsstudierenden ($N = 578$).

²⁷ Für die Vignette mit 6 Items wurden 15 Relationen gebildet, für die Vignette mit 4 Items wurden 6 Relationen gebildet.

welcher Anordnung die übrigen Items gekreuzt wurden. Das ergibt sich daraus, dass jedes Item mit jedem anderen verglichen wird. Analog ist diese Überlegung dazu, wenn der Proband das am schlechtesten geeignete Item erkennt und sein Kreuz bei diesem Item am weitesten links setzt (Anker der Skala am linken Skalenende: 1 = *Trifft überhaupt nicht zu*) (siehe Abbildung 10). Es ist plausibel, dass die beste bzw. die schlechteste Handlungsalternative – im Sinne eines Ankers – für die Probanden am einfachsten zu erkennen ist. Daher ist es relativ einfach mindestens vier korrekte Relationen zu erreichen, da hierzu nur die beste bzw. die schlechteste Handlungsalternative erkannt werden muss. Die erste Punktkategorie „0 Punkte“ soll daher 0 bis 3 richtig gekreuzte Relationen zusammenfassen.

Beispielvignette

Überschrift

Beschreibung der Unterrichtssituation in Form eines kurzen Textes.

Fachdidaktisch relevante Fragestellung

	Trifft überhaupt nicht zu					Trifft voll und ganz zu
1) Der Lehrer sollte ...	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2) Der Lehrer sollte ...	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3) Der Lehrer sollte ...	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4) Der Lehrer sollte ...	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
5) Der Lehrer sollte ...	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Bildung der Itemrelationen und Vergleich mit den Studierendenantworten

	Item 1 mit Item 2	Item 1 mit Item 3	Item 1 mit Item 4	Item 1 mit Item 5	Item 2 mit Item 3	Item 2 mit Item 4	Item 2 mit Item 5	Item 3 mit Item 4	Item 3 mit Item 5	Item 4 mit Item 5
Experten	<	<	<	<	<	<	<	<	<	<
Proband	<	<	<	<	<	<	<	<	<	>

Vergabe der Testpunkte anhand der Itemrelationen

	Item 1 mit Item 2	Item 1 mit Item 3	Item 1 mit Item 4	Item 1 mit Item 5	Item 2 mit Item 3	Item 2 mit Item 4	Item 2 mit Item 5	Item 3 mit Item 4	Item 3 mit Item 5	Item 4 mit Item 5	Summe
Proband	1	1	1	1	1	1	1	1	1	0	9

Abbildung 10. Beispiel einer bearbeiteten Vignette.

In der Abbildung ist beispielhaft das Kreuzmuster eines Probanden dargestellt.

Der Proband in diesem Beispiel hat Item 1 als die (laut Expertenmeinung) am schlechtesten geeignete Handlungsalternative erkannt.

Das Ergebnis „vier korrekte Relationen“ ist wie oben dargestellt in der Regel auf das Erkennen der besten oder der schlechtesten Handlungsalternative zurückzuführen. Um fünf korrekte Relationen zu erreichen, ist exakt eine weitere korrekte Relation notwendig, was relativ unwahrscheinlich ist. Daher wurden vier und fünf korrekte Relationen in die zweite Kategorie „1 Punkt“ zusammengefasst.

Beim nächsten Schritt zur Kategorienbildung wurde das obere Ende der Skala betrachtet: Nur wer alle Relationen korrekt erkennt, erhält die volle Punktzahl. Da dies als ein sehr schwieriges Kriterium erschien (da es auch tatsächlich nur ein einziges Kreuzmuster gibt, das zur vollen Punktzahl führt), sollten neun und zehn korrekte Relationen zur höchsten Kategorie „4 Punkte“ zusammengefasst werden.

Die Kategorienbildung von sechs, sieben und acht korrekten Relationen wurde an den empirischen Ergebnissen zur Verteilung der korrekt erreichten Korrelationen orientiert. Es resultierte folgende Zuordnung von korrekten Relationen zu Testpunkten²⁸:

0 – 3 korrekte Relationen	0 Testpunkte
4 – 5 korrekte Relationen	1 Testpunkt
6 – 7 korrekte Relationen	2 Testpunkte
8 korrekte Relationen	3 Testpunkte ²⁹
9 – 10 korrekte Relationen	4 Testpunkte

Im Vignettentest können daher maximal $12 \times 4 = 48$ Testpunkte erreicht werden. Die einzelnen Vignetten wurden im Anschluss deskriptiv statistisch untersucht. In Tabelle 12 sind die absoluten Häufigkeiten der erreichten Testpunkte für jede Vignette sowie die Mittelwerte und die Standardabweichungen für die einzelnen Vignetten dargestellt.

²⁸ Die Punktvergabe für Vignette 3 (mit 6 Items) und Vignette 7 (mit 4 Items) wurde entsprechend angepasst, sodass auch für diese Vignetten Testpunkte von 0 bis 4 resultierten.

²⁹ Dass für die Vergabe von 3 Testpunkten nur eine Relationenkategorie (nämlich 8 korrekte Relationen) verwendet wurde, hängt mit der empirischen Verteilung der korrekt gebildeten Relationen über alle Vignetten hinweg zusammen; durch die oben dargestellte Kategoriebildung ergibt sich für die Vignetten annähernd eine Normalverteilung für die erreichten Punktwerte (siehe Tabelle 12).

Tabelle 12. Deskriptive Statistiken für die einzelnen Vignetten

	Punktzahl					<i>M</i>	<i>SD</i>	<i>N</i>
	0	1	2	3	4			
Vignette 1	19	89	365	84	21	1.99	0.76	578
Vignette 2	28	209	232	90	19	1.76	0.89	578
Vignette 3	18	46	239	221	54	2.43	0.88	578
Vignette 4	33	104	283	99	59	2.08	0.99	578
Vignette 5	8	114	234	122	100	2.33	1.02	578
Vignette 6	45	106	281	92	54	2.01	1.01	578
Vignette 7	35	118	333	86	6	1.84	0.78	578
Vignette 8	26	106	256	144	36	2.13	0.96	578
Vignette 9	22	67	241	151	97	2.40	1.02	578
Vignette 10	40	107	270	108	53	2.05	1.01	578
Vignette 11	42	107	261	121	47	2.04	1.01	578
Vignette 12	22	108	293	100	55	2.10	0.94	578

Anmerkung. Dargestellt sind die absoluten Anzahlen an erreichten Testpunkten pro Vignette sowie der Mittelwert und die Standardabweichung.

Zuletzt sollte überprüft werden, ob durch die Kategorienbildung möglicherweise wichtige Informationen verloren gehen: Daher wurde der ursprüngliche Testscore mit 0 bis 10 Punkten pro Vignette (da pro Vignette 10 Relationen gebildet werden können) mit dem neuen Testscore mit 0 bis 4 Punkten pro Vignette korreliert. Das Ergebnis von $r = 0.97$ ($p \leq 0.001$) weist darauf hin, dass durch die Kategorienbildung kein bedeutsamer Informationsverlust verursacht wurde.

11.2 Forschungsanliegen 1

Zunächst sollte die Datenstruktur des Vignettentests überprüft werden. Hier ist vor allem von Bedeutung, ob sich die theoretisch angenommene zweidimensionale Struktur des lesedidaktischen Wissens empirisch abbilden lässt. Anschließend soll der Vignettentest mit dem Partial Credit Modell skaliert werden.

11.2.1 Überprüfung der Datenstruktur

Die Datenstruktur wurde konfirmatorisch mit *Mplus* überprüft (Muthén & Muthén, 2012). Hierzu wurden die Daten der Deutschlehramtsstudierenden herangezogen ($N = 578$). Da die Annahme der multivariaten Normalverteilung verletzt ist, wird zur Parameterschätzung der ULSMV³⁰-Schätzer verwendet (Backhaus et al., 2015; Muthén & Muthén, 2012). Da innerhalb der Vignetten aufgrund der Kategorienbildung von 0 bis 4 Testpunkten nicht von intervallskalierten Daten ausgegangen werden kann, wird die Funktion *categorical* für ordinale Variablen verwendet.

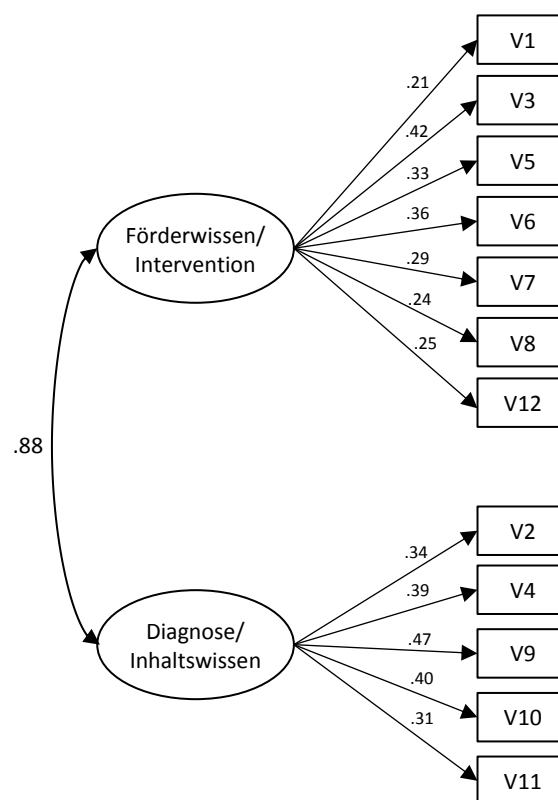


Abbildung 11. Zweidimensionales Modell des lesedidaktischen Wissens mit den Facetten Förderwissen/Intervention und Diagnose/Inhaltswissen.

χ^2 ($N = 578$, $df = 53$) = 71.918, $p = 0.043$, $\chi^2/df = 1.36$, RMSEA = 0.025, CFI = 0.936.

Das zweidimensionale Modell mit den beiden theoretisch angenommenen Facetten „Förderwissen/Intervention“ und „Diagnose/Inhaltswissen“ weist einen guten Fit auf: χ^2 ($N = 578$, $df = 53$) = 71.918, $p = 0.043$, $\chi^2/df = 1.36$, RMSEA = 0.025,

³⁰ ULSMV ist die Abkürzung für *unweighted least squares parameter estimates with standard errors and a mean- and variance-adjusted chi-square test statistic that use a full weight matrix* (Muthén & Muthén, 2012, S. 603).

CFI = 0.936. Die Ladungen der einzelnen Vignetten auf die beiden Facetten fallen gering bis mäßig aus (zwischen 0.21 und 0.47, alle $ps \leq 0.05$). Auffällig ist, dass die beiden Faktoren hoch miteinander korrelieren ($r = 0.88$, $p \leq 0.05$). Dieses Ergebnis weist auf einen starken statistischen wie inhaltlichen Zusammenhang zwischen den beiden Facetten hin. Daher wurde ein alternatives, eindimensionales Modell gerechnet.

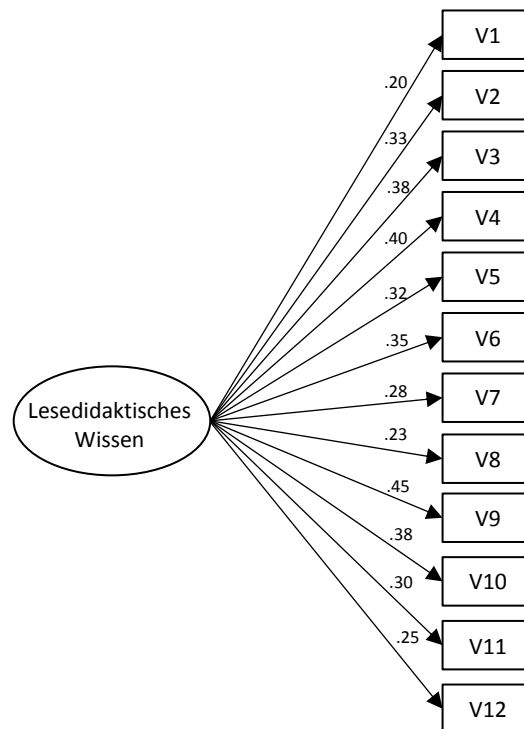


Abbildung 12. Eindimensionales Modell mit dem Faktor „Lesedidaktisches Wissen“. χ^2 ($N = 578$, $df = 54$) = 73.673, $p = 0.039$, $\chi^2/df = 1.36$, RMSEA = 0.025, CFI = 0.933.

Das alternative, eindimensionale Modell weist ebenfalls einen guten Modellfit auf (χ^2 ($N = 578$, $df = 54$) = 73.673, $p = 0.039$, $\chi^2/df = 1.36$, RMSEA = 0.025, CFI = 0.933). Die Ladungen der Vignetten sind vergleichbar hoch (zwischen 0.20 und 0.45, alle $ps \leq 0.05$).

Um zu überprüfen, ob das zweidimensionale Modell des lesedidaktischen Wissens die Daten signifikant besser beschreiben kann als das eindimensionale Modell, wurde ein χ^2 -Differenztest berechnet. Hierzu wurde die *difftest*-Option in *Mplus* verwendet. Für Modelle, die aufgrund einer Verletzung der multivariaten Nor-

malverteilung mit dem ULSMV-Schätzer gerechnet wurden, empfehlen Muthén und Muthén (2012) explizit die Verwendung der *difftest*-Option.

Das Ergebnis des χ^2 -Differenztests lässt darauf schließen, dass das zweidimensionale Modell des lesedidaktischen Wissens die Daten nicht signifikant besser beschreiben kann als das eindimensionale Modell ($\Delta\chi^2 (\Delta df=1) = 1.829, p = 0.176$). Da dieses Ergebnis für eine starke inhaltliche Assoziation der beiden Modellfacetten spricht, wird für die nachfolgenden Analysen von einem eindimensionalen Modell des lesedidaktischen Wissens ausgegangen.

11.2.2 Skalierung

Da die Ergebnisse der konfirmatorischen Faktorenanalyse für ein eindimensionales Modell sprechen, wurde der Vignettentest mit dem Partial Credit Modell eindimensional skaliert. Um die Passung der einzelnen Vignetten zum Modell zu untersuchen, wurde auf Vignettenebene die Infit- und Outfit-Statistik berechnet. Die Werte weisen ausnahmslos auf eine sehr gute Passung der Vignetten zum Modell hin (Werte zwischen 0.98 und 1.00, alle $ps > 0.05$). Der globale Fit weist ebenfalls sehr gute Werte auf (Infit: $M = 1.00, SD = 0.00$; Outfit: $M = 1.00, SD = 0.00$).

Zur Überprüfung von lokaler Abhängigkeit wurde die Q3-Statistik eingesetzt (Yen, 1993). Es wurde zunächst die lokale Abhängigkeit zwischen den gebildeten Itemrelationen innerhalb der Vignetten untersucht. Abbildung 13 zeigt, dass sich zwölf dunkel gefärbte Cluster bilden, die die zwölf Vignetten des Testinstruments repräsentieren. Die dunkle Färbung zeigt einen Q3-Wert > 0.2 an und indiziert damit lokale Abhängigkeit. Die Vermutung, dass zwischen den Items innerhalb der Vignetten lokale Abhängigkeit vorliegt, kann daher bestätigt werden. Dieses Ergebnis bestätigt die Vermutung, dass für den Vignettentest von einem Testlet-design auszugehen ist.

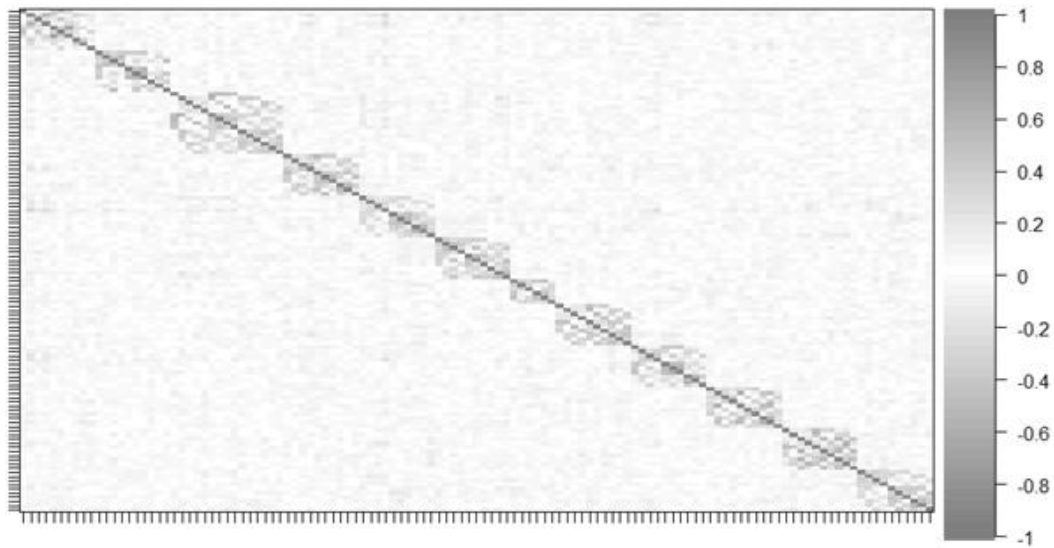


Abbildung 13. Graphische Darstellung der Q3-Statistiken zwischen den Itemrelationen.

Im nächsten Analyseschritt wurde die Q3-Statistik zur Untersuchung der lokalen Abhängigkeit zwischen den Vignetten eingesetzt (siehe Abbildung 14). Hier ist, wie erwartet, keine nennenswerte lokale stochastische Abhängigkeit zu beobachten (alle Werte zwischen -0.03 und 0.19).

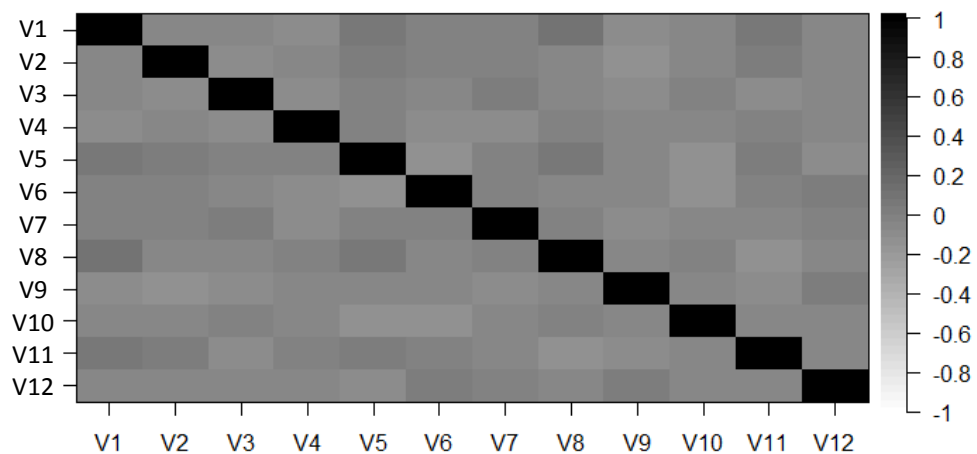


Abbildung 14. Graphische Darstellung der Q3-Statistiken zwischen den Vignetten.

Im nächsten Analyseschritt wurde eine Person-Item Map erstellt. Bei der Skalierung wird von dem Antwortverhalten der Probanden sowohl auf die Itemschwierigkeit als auch auf die Personenfähigkeit der Probanden geschlossen. Hieraus ergibt sich ein wesentlicher Vorteil der Item Response Theorie (IRT) gegenüber der Klassischen Testtheorie (KTT), da die Schätzung der Personenfähigkeit³¹ unabhängig von der Itemschwierigkeit erfolgt. Diese unabhängige Schätzung der beiden latenten Größen Itemschwierigkeit und Personenfähigkeit drückt sich auch in der Bezeichnung Person-Item Map aus: Hier werden sowohl Eigenschaften der Items (Itemschwierigkeit) als auch der Personen (Personenfähigkeit) dargestellt.

Im dichotomen Raschmodell wird die Itemschwierigkeit σ für ein Item als die Stelle der Item-Charakteristischen Funktion (englisch: *item characteristic curve*, ICC) definiert, an der die Lösungswahrscheinlichkeit 50 % beträgt (Wendepunkt der Item-Charakteristischen Funktion). Im Partial Credit Modell hingegen wird für jeden Kategorieübergang ein Schwellenparameter δ_k definiert, der den Übergang von einer Itemkategorie in die nächst höhere beschreibt. δ_k kann nicht wie im dichotomen Raschmodell als „die“ Itemschwierigkeit für ein Item interpretiert werden, sondern beschreibt nur den jeweiligen Übergang von einer Itemkategorie in die nächst höhere. Um im Partial Credit Modell einen allgemeinen Itemschwierigkeitsindex zu erhalten, kann δ berechnet werden: $\delta = \frac{\sum_{k=1}^{m_i} \delta_k}{m}$ (Übersicht: Wu & Adams, 2007).

Im vorliegenden Fall werden für jede Vignette 0 bis 4 Testpunkte vergeben, d. h. pro Vignette werden fünf Fähigkeitsausprägungen abgebildet. Im Partial Credit Modell wird für jeden Kategorieübergang ein Schwellenparameter bestimmt; Schwellenparameter 1 bezieht sich auf den Übergang von Kategorie „0 Testpunkte“ zur Kategorie „1 Testpunkt“ und indiziert das Ausmaß an Personenfähigkeit, das benötigt wird, um die nächsthöhere Kategorie zu erreichen. Schwellenparameter 2 bezieht sich auf den Übergang von Kategorie „1 Testpunkt“ zur Kategorie „2 Testpunkte“, usw..

Im oberen Teil der Person-Item Map (siehe Abbildung 15) ist die Verteilung der Personenfähigkeit abgetragen. Es kann festgestellt werden, dass die Verteilung –

³¹ Es finden sich in der Literatur zu IRT-Methoden verschiedene synonyme Begriffe für das Maß der Personenfähigkeit, u. a. Personenparameter oder Factorscore.

bis auf einige Ausreißer nach oben und unten – annähernd normalverteilt ist. Insgesamt befindet sich die Verteilung leicht rechts vom mittleren Personenparameter (Wert 0). Das bedeutet, dass die befragte Stichprobe im Mittel leicht überdurchschnittlich abschneidet.

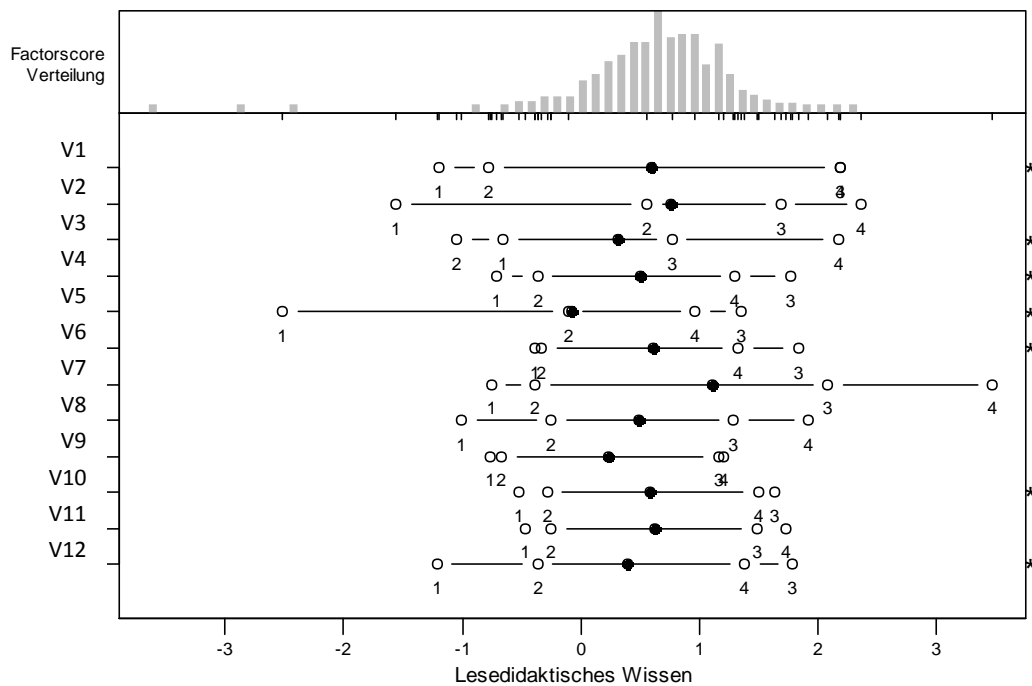


Abbildung 15. Person-Item Map zur Darstellung der Verteilungen der Itemkategorie-schwellen δ_k und der Personenfähigkeiten.

„V“ kennzeichnet die einzelnen Vignetten.

Der schwarze Punkt markiert die mittlere Itemschwierigkeit δ .

Im oberen Bereich der Abbildung ist die Personenfähigkeit abgetragen.

Im unteren Bereich der Abbildung sind die Schwellenparameter δ_k für die einzelnen Vignetten abgetragen. Die Schwellenübergänge beziehen sich, wie oben erläutert, jeweils auf die Kategoriewechsel zur nächst höheren Ausprägung. Es kann beobachtet werden, dass die Schwellenparameter zum Teil vertauscht sind. Dieses Phänomen kann u. a. auftreten, wenn mittlere Itemkategorien empirisch selten besetzt sind. Die Verteilung der absoluten Anzahlen pro Antwortkategorie ist in Tabelle 12 dargestellt.

Weiterhin ist zu beobachten, dass die Schwellen innerhalb der einzelnen Vignetten unterschiedliche Abstände aufweisen. Insgesamt sind die Vignetten alle von relativ gleicher, mittlerer Schwierigkeit.

In der Wright Map sind die kumulierten Lösungswahrscheinlichkeiten (sog. Thurstonian Thresholds) der einzelnen Vignetten dargestellt (Wu & Adams, 2007, S. 50). Diese geben die Stelle der Item-Charakteristischen Funktion an, an der die Wahrscheinlichkeit für eine Versuchsperson 50% oder höher beträgt, eine bestimmte Kategorie oder die nächst höhere zu erreichen. Die kumulierten Lösungswahrscheinlichkeiten weisen per Definition geordnete Schwellen auf. Mit Hilfe der Thurstonian Thresholds kann beurteilt werden, wie schwierig es für Probanden ist, von einer Itemkategorie in die nächsthöhere zu kommen. Verhältnismäßig große Abstände zwischen den Schwellen in Abbildung 16 sprechen dafür, dass ein deutlich höheres Ausmaß an Fähigkeit bei den Probanden vorhanden sein muss, um die nächsthöhere Kategorie zu erreichen. Verhältnismäßig kleine Abstände zwischen den Schwellen deuten darauf hin, dass die Probanden kein deutlich höheres Ausmaß an Fähigkeit benötigen, um die nächsthöhere Kategorie zu erreichen. Je enger die Thurstonian Thresholds beieinander liegen, desto wahrscheinlicher wird das Auftreten vertauschter Itemkategorieschwellen.

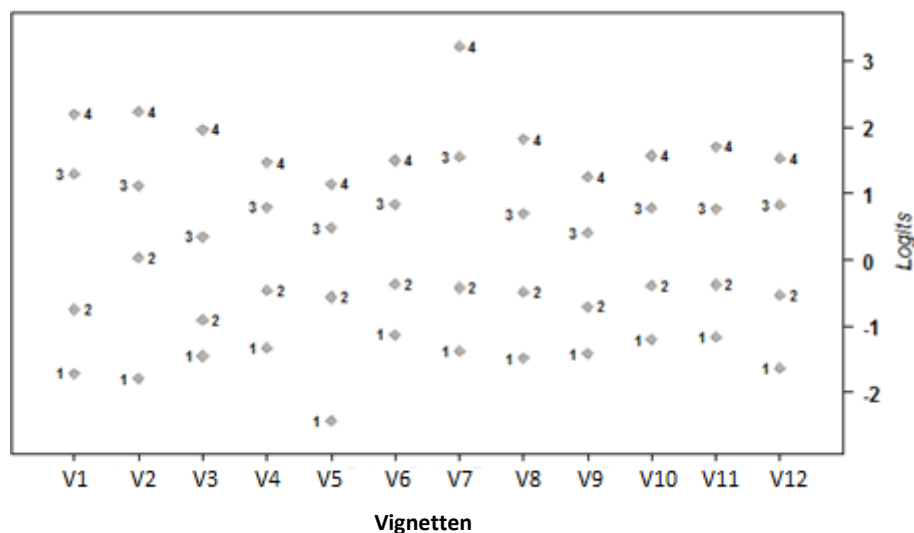


Abbildung 16. Wright Map zur Darstellung der Thurstonian Thresholds. Rechts in der Darstellung sind die Logits³² abgetragen.

Zuletzt wurde die EAP-Reliabilität des Vignettentests bestimmt. Diese fällt mit einem Wert von 0.59 mittelmäßig aus.

³² Die Personen- und Itemparameter werden in der Item Response Theorie in der gleichen Einheit geschätzt, der sog. Logiteinheit (für einen Überblick siehe: Bühner, 2011, S. 495 ff.).

11.3 Forschungsanliegen 2

Im Anschluss an die Analysen zur Datenstruktur und die Skalierung sollten Untersuchungen zur Validität des Testverfahrens durchgeführt werden. Von besonderem Interesse waren hier der Kontrollgruppenvergleich, die Untersuchung des querschnittlichen Anstiegs des lesedidaktischen Wissens über das Studium hinweg sowie der Einfluss der praktischen Erfahrungen der Studierenden auf das lesedidaktische Wissen.

11.3.1 Hypothese 1: Kontrollgruppenvergleich

Zur Überprüfung der Validität des Vignettentests wurde ein Kontrollgruppenvergleich durchgeführt³³. Zunächst wurde überprüft, ob sich die verschiedenen Substichproben hinsichtlich der Abiturdurchschnittsnote und der Leistung bezüglich der Skala zum Sprachlichen Denken unterscheiden³⁴. Dies wurde vermutet, da sich die unterschiedlichen Studiengänge, die im Kontrollgruppenvergleich miteinander verglichen werden, hinsichtlich der Zulassungsvoraussetzungen unterscheiden. Zudem haben andere Studien bereits festgestellt, dass Studierende nicht-gymnasialer Lehrämter schlechtere Schulleistungen zeigen als Studierenden des Gymnasiallehramts (Pohlmann & Möller, 2010). Die Gruppen (Deutschstudierende vs. Studierende fachfremder Lehrämter vs. Psychologiestudierende vs. Germanistik-Bachelor-Studierende) unterscheiden sich sowohl hinsichtlich der Abiturdurchschnittsnote ($F(6) = 83.435$, $p < 0.01$, $\eta_p^2 = 0.34$ ³⁵) als auch im Sprachlichen Denken ($F(6) = 9.855$, $p < 0.001$, $\eta_p^2 = 0.07$) signifikant. Da es sich bei dem Vignettentest um einen Leistungstest handelt, werden für weiteren Analysen die Unterschiede bezüglich der Abiturdurchschnittsnote und des Sprachlichen Denkens – als Indikatoren für allgemeine kognitiven Fähigkeiten – zwischen den Stichproben kontrolliert.

³³ Für die Probanden aus den Kontrollgruppen lagen keine Schätzungen der Personenfähigkeiten vor, da die Daten der Kontrollgruppen nicht skaliert wurden. Aufgrund dessen wurde bei den Analysen bezüglich der Kontrollgruppen der aufsummierte Testscore verwendet. Dieser Testscore weist einen sehr hohen Zusammenhang mit den Personenparametern auf ($r = 0.973$, $p < 0.001$).

³⁴ Im Folgenden wird aus Gründen der Lesbarkeit bei der Skala zum sprachlichen Denken kurz von „Sprachlichem Denken“ gesprochen.

³⁵ Nach Cohen (1988) gibt $\eta_p^2 = 0.1$ einen kleinen, $\eta_p^2 = 0.6$ einen mittleren und $\eta_p^2 > 1.4$ einen starken Effekt an.

Eine einfaktorielle ANCOVA mit den Kovariaten Abiturdurchschnittsnote und Sprachliches Denken zeigt, dass die Deutschlehramtsstudierenden im Vignetten-test besser abschneiden als die Kontrollgruppen. Der geplante einfache Kontrast zwischen den Deutschstudierenden und den Psychologiestudierenden und den Deutschstudierenden und den Germanistikstudierenden erreicht die statistische Signifikanz ($p \leq 0.05$). Der einfache geplante Kontrast zwischen den Deutschstudierenden und den Studierenden fachfremder Lehrämter wird marginal signifikant ($p = 0.07$).

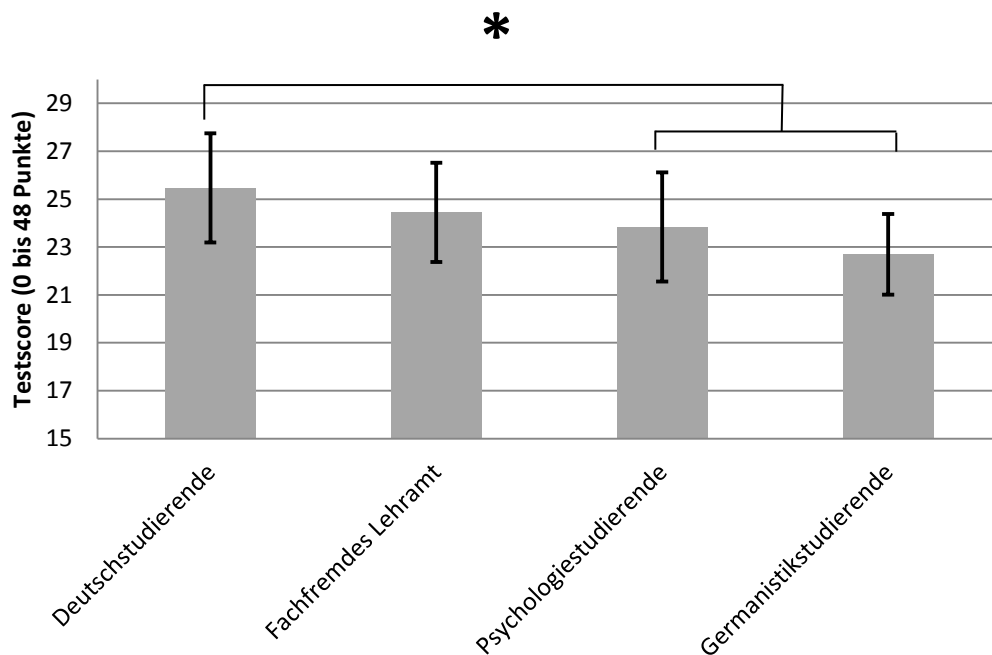


Abbildung 17. Einfaktorielle ANCOVA zum Vergleich zwischen den Kontrollgruppen mit den Kovariaten Abiturdurchschnitt ($p \leq 0.05$) und sprachliches Denken ($p \leq 0.05$), $\eta_p^2 = 0.023$, $*p \leq 0.05$.

11.3.2 Hypothese 2: Anstieg des lesedidaktischen Wissens im Querschnitt

Zunächst wurde untersucht, ob die Studiendauer der Deutschlehramtsstudierenden (operationalisiert über die Semesteranzahl) deren lesedidaktisches Wissen vorher-sagen kann. Hierzu wurde mit *Mplus* ein latentes Regressionsmodell berechnet.

Das Modell weist einen akzeptablen Modellfit auf ($\chi^2 (62) = 96.337$, $p = 0.007$, $\chi^2/df = 1.55$, RMSEA = 0.029, CFI = 0.906). Die Studiendauer kann das lesedi-daktische Wissen der Deutschlehramtsstudierenden mit einem Regressionsge-

wicht von $\beta = 0.06$ vorhersagen ($p \leq 0.05$). Die Residualvarianz des latenten Faktors lesedidaktisches Wissen ist mit dem Wert von 0.98 sehr hoch. Dieses Ergebnis lässt darauf schließen, dass die Studiendauer ein wenig aussagekräftiger Prädiktor des lesedidaktischen Wissens ist.

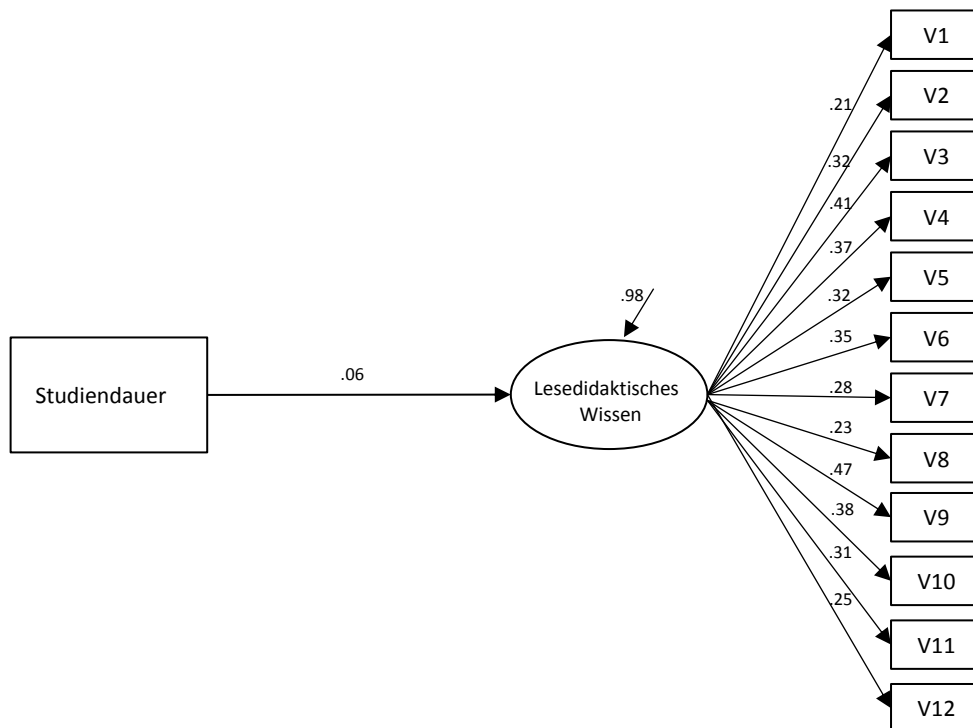


Abbildung 18. Regression des latenten Faktor Lesedidaktisches Wissen auf die Studiendauer (Semester 1 bis Referendariat).

$\chi^2 (62) = 96.337, p = 0.007, \chi^2/df = 1.55, RMSEA = 0.029, CFI = 0.906.$

In einem weiteren Analyseschritt sollten die verschiedenen Substichproben (Primarstufe, Sekundarstufe, Sonderpädagogik, Gymnasiallehramt, fachfremde Lehrämter, Psychologie, Germanistik) auf einen (querschnittlichen) Anstieg des lesedidaktischen Wissens über die Studiendauer hinweg untersucht werden. Hier wurde geprüft, ob sich innerhalb der verschiedenen Substichproben von Beginn bis zum Ende des Studiums ein Anstieg des lesedidaktischen Wissens abbilden lässt.

Hierzu wurden innerhalb der Substichproben zwei Gruppen („Beginn Studium“ und „Ende Studium“) gebildet, die miteinander verglichen wurden³⁶. Deskriptiv kann beobachtet werden, dass in den Deutschlehramtsstudiengängen eine Zunahme des lesedidaktischen Wissens festgestellt werden kann (Säulenpaare 1 bis 4 in Abbildung 19), was für die Kontrollgruppen nicht der Fall ist. Mithilfe von *t*-Tests wurde für die einzelnen Substichproben untersucht, ob der Anstieg des lesedidaktischen Wissens statistisch bedeutsam ist. Dies ist nur für die Substichprobe der Sekundarstufenstudierenden an Pädagogischen Hochschulen der Fall ($t(140) = -2.063, p \leq 0.05, d = 0.22$ ³⁷).

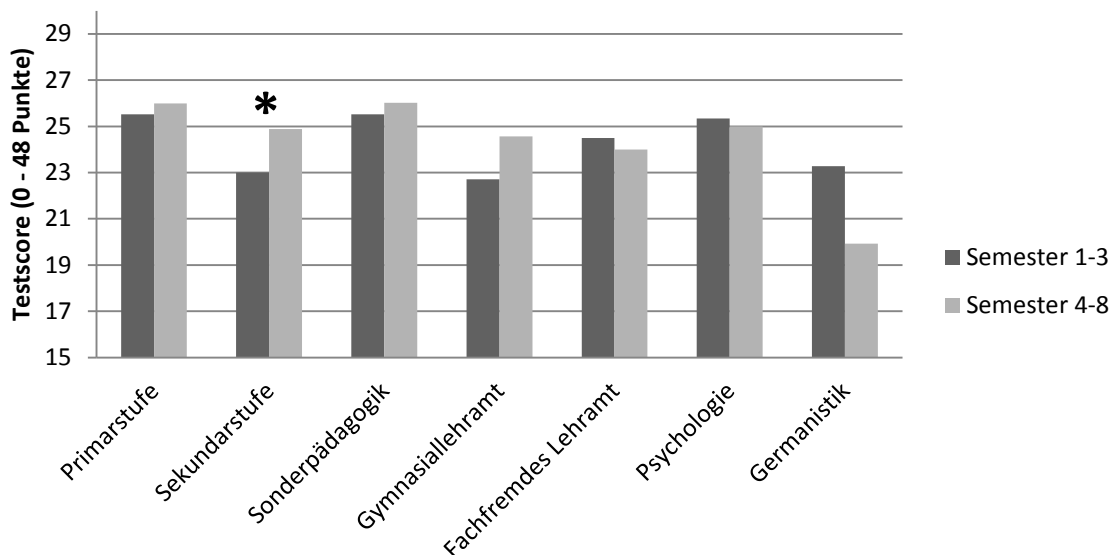


Abbildung 19. *t*-Tests für die verschiedenen Substichproben mit den Gruppen „Beginn des Studiums“ (Semester 1-3) und „Ende des Studiums“ (Semester 4-8), * $p \leq 0.05$.

Da für die Sekundarstufenstudierenden an den Pädagogischen Hochschulen im Querschnitt ein signifikanter „Anstieg“ des lesedidaktischen Wissens von Beginn bis Ende des Studiums verzeichnet werden kann, wird dieser „Verlauf“ genauer untersucht und dargestellt.

³⁶ Die Bildung der beiden Gruppen „Beginn des Studiums“ (Semester 1-3) und „Ende des Studiums“ (Semester 5-8) ergab sich daraus, dass die Stichprobengrößen in manchen Gruppen eingeschränkt war. Durch diese Aufteilung wurde ein sinnvoller Vergleich mit mind. 15 Versuchspersonen pro Zelle möglich.

³⁷ Nach Cohen (1988) gibt $d = 0.2$ einen kleinen, $d = 0.5$ einen mittleren und $d = 0.8$ einen starken Effekt an.

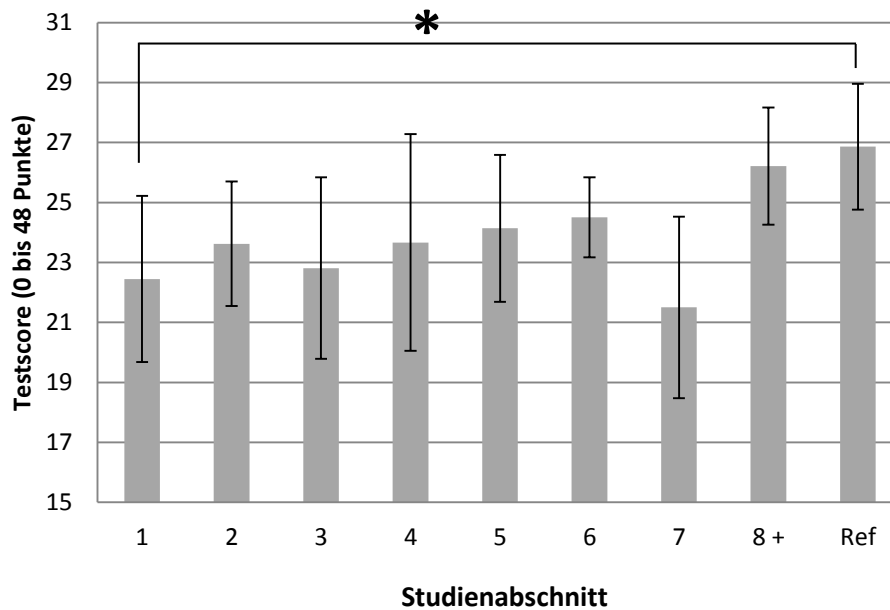


Abbildung 20. Einfaktorielle ANOVA (n. s.) zum Anstieg des lesedidaktischen Wissens über die Semester für die PH-Sekundarstufendeutschlehramtsstudierenden mit einfachen geplanten Kontrasten zum ersten Semester, * $p \leq 0.05$. Die Fehlerbalken indizieren einfache Standardabweichungen. „8+“ bezeichnet Studierende, die im 8. oder einem höheren Semester studieren, „Ref“ bezeichnet Studierende im Referendariat.

Im Querschnitt ist ein kontinuierlicher Anstieg des lesedidaktischen Wissens zu beobachten, wobei der einfache geplante Kontrast zwischen dem ersten Semester und dem Referendariat signifikant wird. Hier ergibt sich ein starker Effekt von $d = 0.81$. In Semester 7 fällt der Einbruch des lesedidaktischen Wissens auf; allerdings ist dieser nicht (zuverlässig) interpretierbar, da in dieser Gruppe nur 4 Probanden befragt wurden (siehe Tabelle 11).

11.3.3 Hypothese 3: Einfluss der Lehrerfahrung auf das lesedidaktische Wissen

Es sollte die Annahme überprüft werden, ob die bereits erworbene Lehrerfahrung der Studierenden (allgemein bzw. spezifisch im Leseunterricht) die Testleistung im Vignettentest vorhersagen kann.

Hierzu wurden mit *Mplus* latente Regressionen des lesedidaktischen Wissens auf die Lehrerfahrung im Allgemeinen sowie die Lehrerfahrung im Leseunterricht berechnet. Da die Variablen zur Lehrerfahrung in einem dichotomen Format vorliegen („ja“/„nein“) wurde für diese Variablen die Funktion *categorical* verwendet.

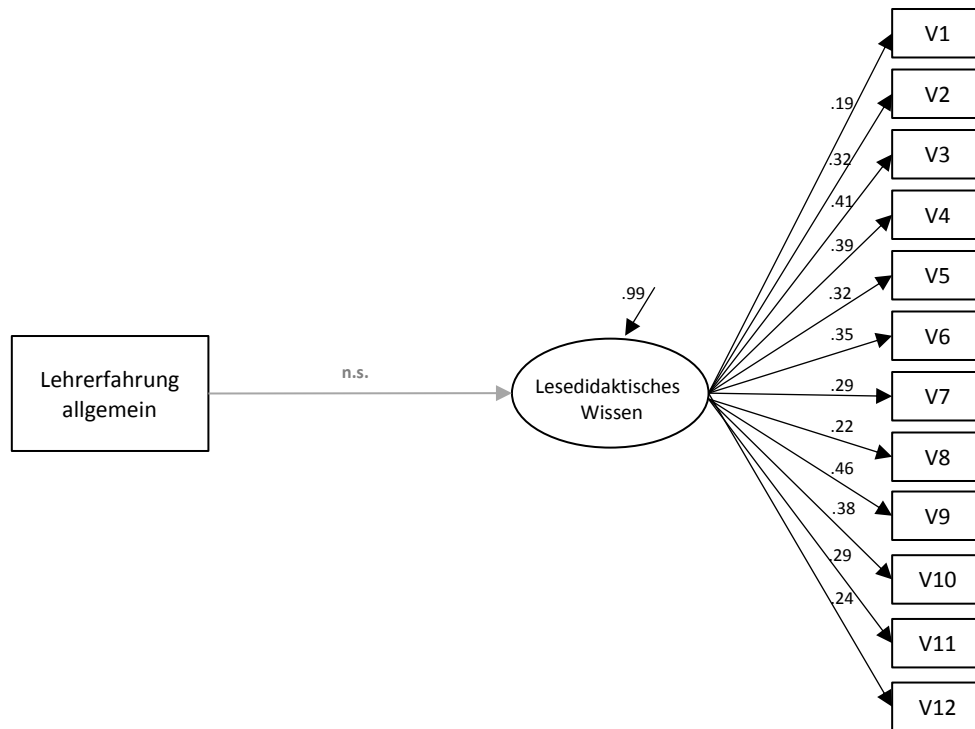


Abbildung 21. Regression des latenten Faktor Lesedidaktisches Wissen auf die Lehrerfahrung im Allgemeinen (ja/nein).

$\chi^2 (65) = 74.105, p = 0.206, \chi^2/df = 1.14, RMSEA = 0.016, CFI = 0.950.$

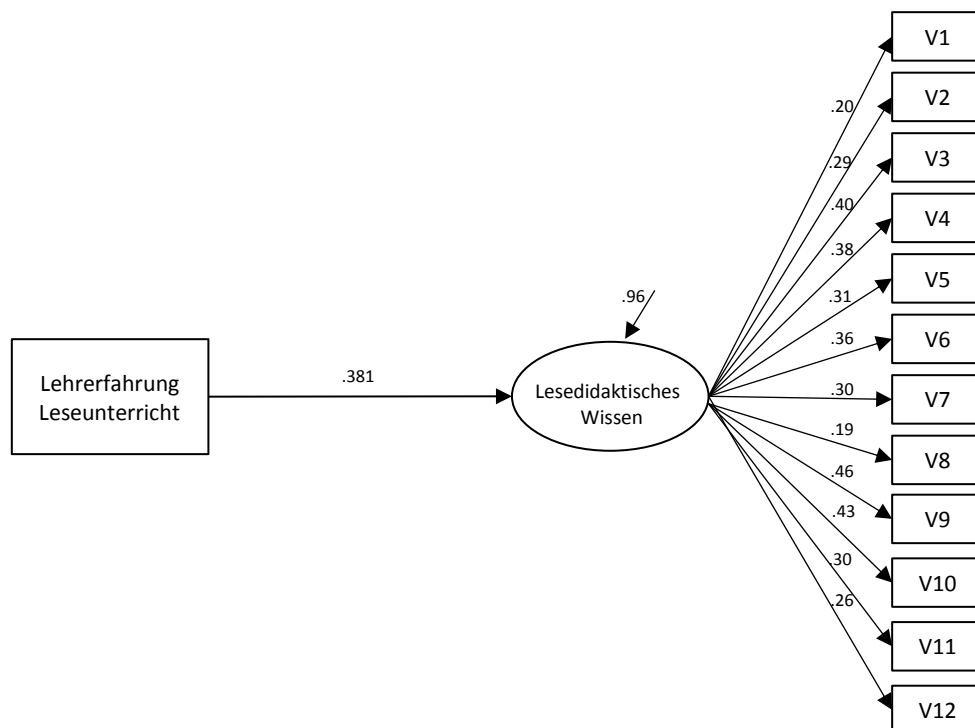


Abbildung 22. Regression des latenten Faktor Lesedidaktisches Wissen auf die Lehrerfahrung im Leseunterricht (ja/nein).

$\chi^2 (62) = 78.413, p = 0.122, \chi^2/df = 1.26, RMSEA = 0.019, CFI = 0.930.$

Es zeigt sich, dass die Lehrerfahrung im Leseunterricht das lesedidaktische Wissen der Studierenden mit einem Regressionsgewicht von $\beta = 0.381$ ($p \leq 0.05$) vorhersagen kann. Die Lehrerfahrung im Allgemeinen ist kein signifikanter Prädiktor für das lesedidaktische Wissen.

11.3.4 Hypothese 4: Einfluss lesedidaktischer Seminare auf das lesedidaktische Wissen

Es wurde erwartet, dass der Besuch lesedidaktischer Seminare die Testleistung vorhersagen kann. Um diese Annahme zu überprüfen, wurde erneut in *Mplus* eine latente Regression des lesedidaktischen Wissens auf den Besuch lesedidaktischer Seminare berechnet (Muthén & Muthén, 2012). Da die Variable „Seminarbesuch“ in einem dichotomen Format vorliegt („ja“/„nein“) wurde für diese Variable die Funktion *categorical* verwendet.

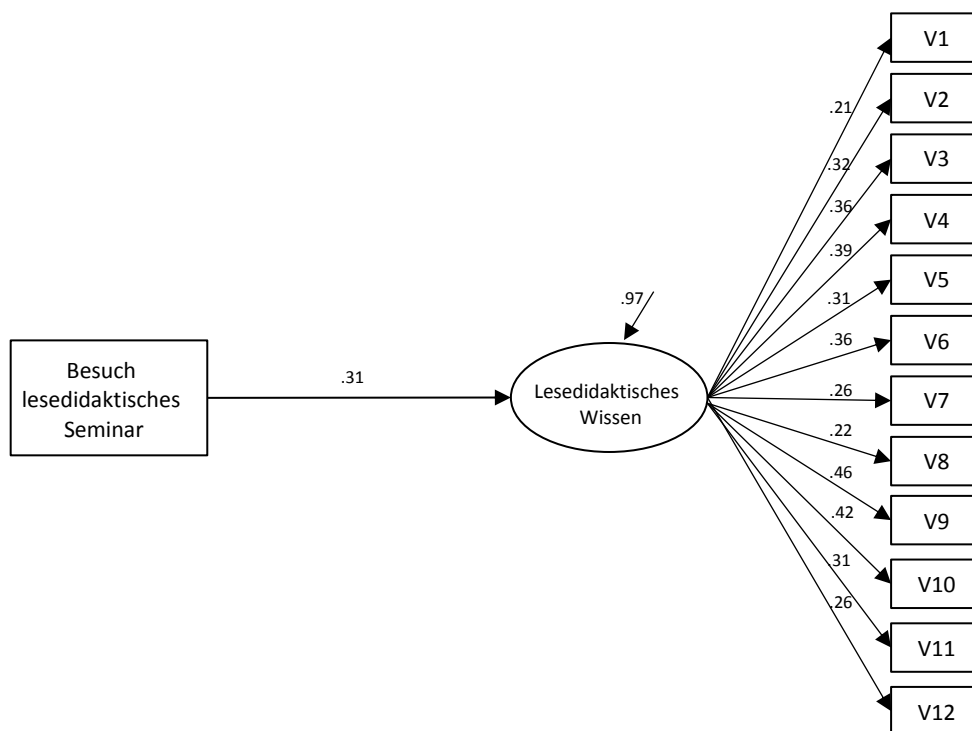


Abbildung 23. Regression des latenten Faktor Lesedidaktisches Wissen auf „Besuch lesedidaktisches Seminar“ (ja/nein).

$\chi^2(62) = 74.175$, $p = 0.204$, $\chi^2/df = 1.19$, RMSEA = 0.016, CFI = 0.952.

Der Besuch eines lesedidaktischen Seminars kann das lesedidaktische Wissen der Studierenden mit einem Regressionsgewicht von $\beta = 0.31$ ($p \leq 0.05$) vorhersagen.

11.3.5 Hypothese 5: Praktische Erfahrung und Studiendauer als inkrementelle Prädiktoren des lesedidaktischen Wissens

In einem weiterführenden Analyseschritt sollte untersucht werden, inwiefern die praktischen Erfahrungen der Deutschstudierenden die Testleistung über die Studiendauer hinaus vorhersagen können. Für den latenten Faktor Praktische Erfahrungen wurden die beiden Indikatoren Lehrerfahrung im Leseunterricht und Besuch lesedidaktischer Seminare verwendet. Das latente Regressionsmodell weist eine sehr gute Passung auf ($\chi^2 (88) = 100.368$, $p = 0.173$; $\chi^2/df = 1.14$; RMSEA = 0.016; CFI = 0.964).

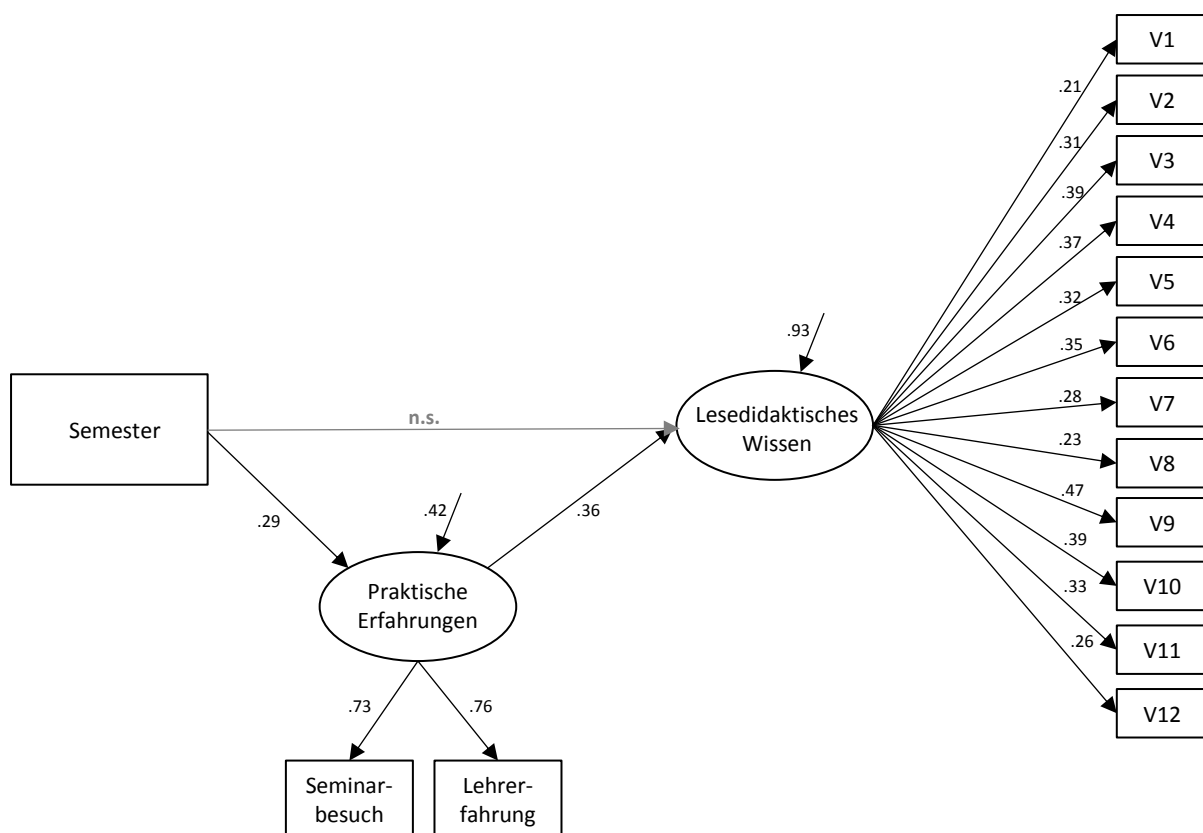


Abbildung 24. Latentes Regressionsmodell für die Regression des lesedidaktischen Wissens auf die Semesteranzahl und den latenten Faktor „Praktische Erfahrungen“.

$\chi^2 (N = 576, df = 88) = 100.368$, $p = 0.173$; $\chi^2/df = 1.14$; RMSEA = 0.016; CFI = 0.964.

Die Semesteranzahl kann die Praktischen Erfahrungen – erwartungsgemäß – mit einem Regressionsgewicht von $\beta = 0.29$ ($p \leq 0.05$) statistisch bedeutsam vorher-sagen, das lesedidaktische Wissen jedoch nicht. Durch den Regressor Semester-

anzahl wird etwa 57 % der Varianz der latenten Variable Praktische Erfahrungen aufgeklärt.

Der latente Faktor Praktische Erfahrungen sagt das lesedidaktische Wissen mit einem Regressionsgewicht von $\beta = 0.36$ ($p \leq 0.05$) vorher. Des Weiteren ist abzulesen, dass die Praktischen Erfahrungen der Studierenden ca. 7 % der Varianz des latenten Faktors des lesedidaktischen Wissens erklärt.

Es handelt sich hier um einen Mediationseffekt: Die Vorhersage des lesedidaktischen Wissens der Deutschlehramtsstudierenden durch die Semesteranzahl wird über den latenten Faktor Praktische Erfahrungen vollständig mediiert (Baron & Kenny, 1986).

11.4 Forschungsanliegen 3

Neben dem Vignettentest wurden in der Studierendenbefragung PK, lesedidaktische Überzeugungen, die Berufswahlmotivation, Aspekte des berufsbezogenen Selbstkonzeptes und die BIG-5 erfasst, um mögliche Zusammenhänge zum lesedidaktischen Wissens (explorativ) zu untersuchen. Hierzu wurden in einem ersten Analyseschritt bivariate Korrelationen zwischen dem Testscore und den erfassten Konstrukten durchgeführt³⁸.

³⁸ Die Analysen zu Forschungsfrage 3 basieren auf $N = 531$ Lehramtsstudierenden (diejenigen, die nicht gymnasiales Lehramt studieren). Bei der Teilstichprobe der Gymnasiallehramtsstudierenden wurden die genannten Konstrukte nicht abgefragt (missing by design).

Tabelle 13. Bivariate Korrelationen zwischen dem Testscore (lesedidaktisches Wissen) und verwandten Konstrukten

Variable	2	3	4	5	6	7	8	9	10
1 Testscore	.15 [*]	-.13 ^{**}	.20 ^{**}	.19 ^{**}	-.15 ^{**}	.20 ^{**}	-.07	.11 [*]	.11 [*]
2 SD		-.16 ^{**}	.15 ^{**}	.14 ^{**}	-.18 ^{**}	.06	.11 [*]	-.01	.05
3 Abi			-.15 ^{**}	-.10 [*]	.10 [*]	-.13 ^{**}	-.07	.01	.01
4 PK				.13 ^{**}	-.15 ^{**}	.19 ^{**}	-.10 [*]	.06	.06
5 KÜ					-.43 ^{**}	-.47 ^{**}	.19 ^{**}	.15 ^{**}	.06
6 DÜ						.00	-.27 ^{**}	.05	.03
7 GÜ							-.11 ^{**}	.28 ^{**}	.11 [*]
8 SÜ								-.11 ^{**}	-.10 [*]
9 PI									.25 ^{**}
10 V									---

Anmerkungen. Testscore = Summenscore (0 – 48 Punkte), SD = Sprachliches Denken, Abi = Abiturdurchschnitt, PK = Pädagogisches Unterrichtswissen, KÜ = Konstruktivistische Überzeugung im Leseunterricht, DÜ = Direkt-transmissive Überzeugung im Leseunterricht, GÜ = Globale Überzeugung im Leseunterricht, SÜ = Überzeugung zu Instruktionsstrategien im Leseunterricht, PI = Pädagogisches Interesse, V = Verträglichkeit.

Aus Gründen der Übersichtlichkeit werden nur die signifikanten Korrelationen zum Testscore dargestellt.

* $p \leq .05$, ** $p \leq .01$.

Der Befund, dass der Testscore mit der Abiturdurchschnittsnote sowie dem Sprachlichen Denken korreliert, weist darauf hin, dass neben dem lesedidaktischen Wissen der Studierenden durch den Vignettentest auch allgemein kognitive Fähigkeiten erfasst werden.

Daher wurde in einem zweiten Analyseschritt partielle Korrelationen zwischen dem Testscore und den verwandten Konstrukten berechnet, wobei die Abiturdurchschnittsnote und Sprachliches Denken als Kovariaten in die Analyse aufgenommen wurden. So soll der – um den Einfluss der Indikatoren für allgemeine kognitive Fähigkeiten bereinigte – Zusammenhang zwischen dem lesedidaktischen Wissen und den erfassten Konstrukten untersucht werden.

Tabelle 14. Partielle Korrelation des Testscores mit weiteren Konstrukten unter Kontrolle von Abiturdurchschnittsnote und Sprachlichem Denken

Variable	2	3	4	5	6	7	8
1 Testscore	.15**	.17**	-.11*	.18**	-.09*	.10*	.10*
2 PK		.10*	-.11*	.16**	-.12**	.05	.05
3 KÜ			-.41**	.15*	.20**	.14**	.04
4 DÜ				.03	-.25**	.05	.04
5 GÜ					-.11**	.27**	.11*
6 SÜ						-.10*	-.12**
7 PI.							.25**
8 V							---

Anmerkungen. Testscore = Summenscore (0 – 48 Punkte), PK = Pädagogisches Unterrichtswissen, KÜ = Konstruktivistische Überzeugung im Leseunterricht, DÜ = Direkttransmissive Überzeugung im Leseunterricht, GÜ = Globale Überzeugung im Leseunterricht, SÜ = Überzeugung zu Instruktionsstrategien im Leseunterricht, PI = Pädagogisches Interesse, V= Verträglichkeit.

Aus Gründen der Übersichtlichkeit werden nur die signifikanten Korrelationen zum Testscore dargestellt.

* $p \leq .05$, ** $p \leq .01$.

Durch die statistische Kontrolle der Indikatoren für allgemeine kognitive Fähigkeiten werden die Zusammenhänge zwischen dem Testscore und den weiteren Konstrukten zum Teil etwas abgeschwächt.

11.4.1 Hypothese 6: Zusammenhang zu PK

In Hypothese 6 wurde ein positiver Zusammenhang zwischen dem lesedidaktischen Wissen und PK postuliert. Dieser Zusammenhang konnte erwartungsgemäß gezeigt werden. Durch die Kontrolle der allgemeinen kognitiven Fähigkeiten der Studierenden sinkt die Korrelation etwas ab (von $r = 0.20$ auf $r = 0.15$, $ps \leq 0.01$). Dieses Ergebnis verstärkt die Annahme, dass durch die Testverfahren zum lesedidaktischen Wissen und zu PK Wissensbereiche der Studierenden erfasst werden, die von allgemeinen kognitiven Fähigkeiten abgegrenzt werden können und die, über die allgemeine kognitive Fähigkeit hinaus, miteinander in Zusammenhang stehen.

11.4.2 Hypothese 7: Zusammenhang zu lesedidaktischen Überzeugungen

Der Zusammenhang zu spezifischen Überzeugungen im Leseunterricht wurde in Hypothese 7 explorativ untersucht. Nach der Kontrolle der Indikatoren für allgemeine kognitive Fähigkeiten kann festgestellt werden, dass das lesedidaktische Wissen einen positiven Zusammenhang zu einer konstruktivistischen ($r = 0.17$, $p \leq 0.01$) und zur globalen ($r = 0.18$, $p \leq 0.01$) Überzeugung im Leseunterricht aufweist. Zur direkt-transmissiven ($r = -0.11$, $p \leq 0.05$) Überzeugung sowie zur Überzeugung zur Vermittlung expliziter Lesestrategien ($r = -0.09$, $p \leq 0.05$) sind negative Zusammenhänge zu beobachten.

11.4.3 Hypothese 8: Zusammenhang zum berufsbezogenen Selbstkonzept

Es wurde erwartet, dass ein positiver Zusammenhang zwischen dem berufsbezogenen Selbstkonzept und dem lesedidaktischen Wissen der Lehramtsstudierenden zu beobachten ist. Entgegen der Erwartungen sind keine statistisch bedeutsamen Zusammenhänge zwischen den Skalen zum berufsbezogenen Selbstkonzept (Selbstkonzept Fach, Selbstkonzept Diagnostik, Selbstkonzept Erziehung) und dem lesedidaktischen Wissen feststellbar.

11.4.4 Hypothese 9: Zusammenhang zur Berufswahlmotivation

Es sollte explorativ untersucht werden, ob intrinsische Aspekte der Berufswahlmotivation positive und extrinsische Aspekte der Berufswahlmotivation negative Zusammenhänge mit dem lesedidaktischen Wissen aufweisen. Ausschließlich die Skala Pädagogisches Interesse erreicht nach der Kontrolle der allgemeinen kognitiven Leitungsfähigkeit die statistische Signifikanz ($r = 0.10$, $p \leq 0.05$). Dieser Zusammenhang ist erwartungsgemäß positiv, da es sich bei pädagogischem Interesse um einen Aspekt der intrinsischen Motivation handelt.

11.4.5 Hypothese 10: Zusammenhang zu den BIG-5

Es wurde erwartet, dass das lesedidaktische Wissen positive Zusammenhänge zu den Persönlichkeitseigenschaften Extraversion und Gewissenhaftigkeit aufweist. Diese Zusammenhänge konnten nicht festgestellt werden. Allerdings korreliert

das lesedidaktische Wissen unerwarteterweise mit der Persönlichkeitseigenschaft Verträglichkeit ($r = 0.25, p \leq 0.01$).

12 Diskussion

Die vorliegende Arbeit thematisiert die Struktur sowie die Entwicklung professioneller Wissensinhalte von Deutschlehramtsstudierenden mit einem Fokus auf das fachdidaktische Wissen im Leseunterricht. Um das lesedidaktische Wissen valide und reliabel abbilden zu können, wurde in dieser Studie ein Vignettentest mit geschlossenem Antwortformat für den Bereich der Lesedidaktik erstellt und eingesetzt. Durch dieses Testformat sollten (auch) die impliziten Anteile des fachdidaktischen Wissens im Leseunterricht erfassbar gemacht werden. Die Ergebnisse zur Inhaltsvalidität, zur Datenstruktur und Skalierung, zur querschnittlichen Entwicklung des lesedidaktischen Wissens über das Studium hinweg, dem Einfluss der praktischen Erfahrungen auf die Entwicklung des lesedidaktischen Wissens sowie zu Zusammenhängen zwischen dem lesedidaktischen Wissen und weiteren Konstrukten werden im Folgenden erläutert und diskutiert.

12.1 Inhaltsvalidität und Auswertungsnorm

Die inhaltliche Validität der Testaufgaben wurde im Prozess der Testentwicklung von Experten aus verschiedenen Bereichen der lesedidaktischen Forschung und Praxis durch ein mehrstufiges Ratingverfahren überprüft. Die Befragung ausgewiesener Experten der jeweiligen Domäne zur Auswahl und Optimierung fachdidaktischer Testaufgaben hat sich bereits in Vorgängerstudien bewährt (König, Blömeke & Schwippert, 2013; Krauss et al., 2011; Tepner & Dollny, 2014). Die zwölf Vignetten des entwickelten Testinstruments wurden durch die Experten durchgängig als fachdidaktisch relevant, eindeutig beschrieben und nahe am Unterrichtsalltag eingeschätzt. Die inhaltliche Validität des Vignettentests kann daher als gegeben angesehen werden.

Im Verlauf der inhaltlichen Validierung der Testaufgaben wurden die Vignetten des Themenbereichs „Aufgabenschwierigkeit“ aus dem Testinstrument ausgeschlossen. Die Ursache hierfür war, dass sich die befragten Experten hinsichtlich der Bewertung der Items nicht ausreichend einig waren. Daher stellt sich die Frage, weshalb gerade für diesen Themenbereich deutlich uneinheitlichere (quantita-

tive) Expertenratings für die Items resultierten als für die anderen Themenbereiche: Es könnte sein, dass hinsichtlich der Beurteilung der Aufgabenschwierigkeit von Leseaufgaben tatsächlich heterogene Expertenmeinungen in der Scientific Community vertreten sind und sich dieser Dissens in den uneinheitlichen Ratings äußert. Andererseits könnte es sein, dass die Vignettenstämme und/oder die zugehörigen Items nicht optimal konstruiert wurden (bspw. bezüglich der Eindeutigkeit der Formulierungen oder des Inhalts), sodass die Diskrepanz zwischen den Expertenratings hierdurch erzeugt worden sein könnte. An dieser Stelle wären erneute qualitative Expertenbefragungen angebracht, um den Ursachen für die Uneinheitlichkeit bei der Bewertung der Items weiter nachzugehen. In der Tat ist die Beurteilung der Aufgabenschwierigkeit von Leseaufgaben ein wichtiges Kompetenzmerkmal von (angehenden) Deutschlehrkräften (vgl. Artelt, 2009) weshalb es wünschenswert wäre, wenn auch inhaltlich valide Vignetten zu diesem Themenbereich im Testinstrument eingesetzt werden könnten.

Die Lösungsmuster für die einzelnen Vignetten wurden durch eine abschließende Expertenbefragung empirisch abgeleitet (Bildung einer sog. Expertennorm) (vgl. Tepner & Dollny, 2014). Um die Güte der Expertennorm zu überprüfen, wurde der Intraklassen-Korrelationskoeffizient (ICC) herangezogen, der eine Aussage hinsichtlich der Übereinstimmung der Expertenantworten zulässt. Eine Expertennorm kann nur dann als zuverlässige Referenz zur Bewertung der Studierendenantworten herangezogen werden, wenn sich die Experten auch weitestgehend einig über die konkrete Beantwortung der Testaufgaben sind. Der ICC-Wert von $r = 0.87$ lässt auf eine zufriedenstellende Einheitlichkeit der Expertenurteile schließen, sodass von einem verlässlichen Lösungsmuster für die Bewertung der Studierendenantworten ausgegangen werden kann (Wirtz & Caspar, 2002).

12.2 Datenstruktur

Bevor die Analysen zur Datenstruktur durchgeführt wurden, sollte überprüft werden, ob die Zugehörigkeit der Studierenden zu einer gewissen Hochschule die Testleistung im Vignettentest beeinflusst. Der dazu berechnete ICC weist mit 0.011 einen sehr geringen Wert auf; das bedeutet, dass gerade einmal 1.1 % der Varianz des lesedidaktischen Wissens auf die unterschiedliche Hochschulzugehörigkeit der befragten Studierenden zurückzuführen ist. Dieses Ergebnis weist da-

rauf hin, dass die Ausprägung des lesedidaktischen Wissens nicht vom Studium an einer bestimmten Hochschule abhängig ist; die Varianzaufklärung liegt somit bei anderen Faktoren als der Hochschulzugehörigkeit.

Auch andere Autoren kommen zu vergleichbaren Ergebnissen: Bezüglich des bildungswissenschaftlichen Wissens von Lehramtsstudierenden konnten ähnliche ICCs festgestellt werden (Kunina-Habenicht et al., 2013). Die Autoren schlussfolgern daraus, dass der Großteil der Varianz in der Testleistung innerhalb der einzelnen Hochschulen liegt und dass der Einfluss der Qualität sowie der individuellen Nutzung der Lerngelegenheiten auf die Entwicklung professioneller Wissensinhalte als enorm angesehen werden kann und muss.

Das theoretisch angenommene, zweidimensionale Modell des lesedidaktischen Wissens mit den beiden Facetten „Förderwissen/Intervention“ und „Diagnose/Inhaltswissen“ wurde konfirmatorisch überprüft. Die globalen Modellfitindices weisen auf eine gute Modellpassung der Daten hin (Hu & Bentler, 1998; Jöreskog & Sörbom, 1993). Allerdings weisen die beiden Facetten einen hohen korrelativen Zusammenhang von $r = 0.87$ auf. Fraglich ist daher, ob hier von zwei inhaltlich eindeutig abgrenzbaren Wissensbereichen des lesedidaktischen Wissens ausgegangen werden sollte. Dieses Ergebnis lässt eher darauf schließen, dass die beiden Wissensbereiche „Förderwissen/Intervention“ und „Diagnose/Inhaltswissen“ eng miteinander verzahnt sind.

Andere Studien berichten folgende Ergebnisse zu Zusammenhängen von fachdidaktischen Wissensbereichen: In der COACTIV-Studie wurde das fachdidaktische Wissen von Mathematiklehrkräften durch die drei Subskalen *Aufgaben*, *Schülerkognitionen* und *Erklären und Repräsentieren* erfasst. Die Ladungen der Subskalen auf den latenten Faktor *Fachdidaktisches Wissen* liegen zwischen 0.56 und 0.74 (Krauss et al., 2011, S. 148). Bezüglich des Testinstruments zum professionellen Wissen im Deutschunterricht der Studie TEDS-LT berichten Bremerich-Vos und Dämmer (2013) eine latente Korrelation von $r = 0.69$ zwischen den beiden fachdidaktischen Kompetenzfacetten Sprachdidaktik und Literaturdidaktik (Bremerich-Vos & Dämmer, 2013, S. 59).

Der Zusammenhang zwischen den beiden Facetten des lesedidaktischen Wissens „Diagnose/Inhaltswissen“ und „Förderwissen/Intervention“ fällt höher aus als Zu-

sammenhänge zwischen Facetten bzw. Inhaltsbereichen des fachdidaktischen Wissens in vergleichbaren Studien. Daher wurde ein alternatives, eindimensionales Modell berechnet. Das eindimensionale Modell des lesedidaktischen Wissens weist ebenfalls einen guten globalen Modellfit auf (Hu & Bentler, 1998; Jöreskog & Sörbom, 1993). Auch die Höhe der Ladungen der Vignetten sind im ein- und zweidimensionalen Modell vergleichbar. Da der χ^2 -Differenztest zeigt, dass das zweidimensionale Modell des lesedidaktischen Wissens die Daten nicht signifikant besser beschreiben kann als das eindimensionale Modell. Daher – sowie nach dem Prinzip der Parsimonität – wird von einem eindimensionalen Modell des lesedidaktischen Wissens ausgegangen, das aus zwei sehr eng miteinander verwobenen Wissensbereichen besteht.

Dieses Ergebnis ist anschlussfähig an Autoren, die argumentieren, dass verschiedene PCK-Komponenten sehr eng miteinander verbunden sind und eine Aufteilung von PCK in inhaltliche Teilbereiche daher wenig sinnvoll erscheint (Cochran, King & DeRuiter, 1991). Fernández-Balboa und Stiehl (1995) betonen sogar die Wichtigkeit der Integration der verschiedenen Komponenten von PCK:

„It is only when the teacher is able to integrate them all [die Komponenten] and apply them appropriately (at the right time, for the right students, in the right circumstances) that superior teaching will occur. In other words, it is not the separate existence, but rather the intersection and rightful integration of all these PCK components that comprises good teaching. Interestingly, due to the interjective and integrative character of PCK, enhancing any of these components will also enhance PCK as a whole.” (Fernández-Balboa & Stiehl, 1995, S. 294).

Bei der Modellspezifikation fallen die geringen bis mäßigen Ladungen der Vignetten auf den Überfaktor „lesedidaktisches Wissen“ auf (zwischen 0.20 und 0.45, alle $ps \leq 0.05$). Dies hängt vermutlich vorrangig damit zusammen, dass die Unterrichtssituationen sehr verschiedene Bereiche des Leseunterrichts thematisieren und somit kein homogenes Konstrukt erfasst wird. Für Testinstrumente mit heterogenen Testaufgaben bzw. heterogenen Item Parcels werden häufig geringe Ladungen berichtet (Cole, Perkins & Zelkowitz, 2015).

Daneben können die geringen Ladungen mit dem (noch) geringen lesedidaktischen Wissen der Lehramtsstudierenden in Zusammenhang stehen. Möglicherweise gelingt es den Lehramtsstudierenden als Novizen ihres Faches (noch) nicht gut, die Unterrichtssituationen in den Vignetten transsituational professionell zu beurteilen. Das würde bedeuten, dass die Studierenden manche Vignetten schon gut – im Sinne der optimalen Lösung des aggregierten Experten – bearbeiten können, andere wiederum nicht. In der Literatur wird passend dazu angenommen, dass PCK von Lehramtsstudierenden sowie Referendaren (noch) nicht als stabiler professioneller Wissensinhalt anzusehen ist, sondern in fragmentierter Form vorliegt (Lee, Brown, Luft & Roehrig, 2007). Ergebnisse aus den Naturwissenschaften sprechen davon, dass PCK bei Lehramtsanwärtern (noch) in rudimentärer Form vorliegt (Aydeniz & Kirbulut, 2014). Eine Festigung des lesedidaktischen Wissens ist mit zunehmender Ausbildungsdauer und praktischer Erfahrung der Studierenden zu erwarten (Friedrichsen et al., 2008), u. a. durch den Erwerb der sog. Prozeduralisierungskompetenz (siehe Kapitel 5.4, Neuweg, 2005; Patry, 2000). Diese Annahme steht in Einklang mit den Ergebnissen dieser Studie, dass das lesedidaktische Wissen der angehenden Lehrkräfte u. a. durch den Erwerb von Lehrerfahrung sowie die Nutzung spezifischer Lerngelegenheiten bedingt wird (siehe Kapitel 5.5).

12.3 Skalierung

Neben der Untersuchung der theoretisch angenommenen Struktur des lesedidaktischen Wissens war die Analyse der psychometrischen Eigenschaften der Testaufgaben eine zentrale Fragestellung, die im Rahmen dieser Dissertation bearbeitet wurde. Um differenzierte Kenntnisse über die Testeigenschaften zu gewinnen, wurde der Vignettentest mit einem Partial Credit Modell skaliert. Hinsichtlich der Infit- und Outfit-Statistiken können sehr gute Fit-Indices beobachtet werden (Adams & Wu, 2002; Ayala, 2009).

Die gebildeten Itemrelationen innerhalb der einzelnen Vignetten weisen deutlich lokale stochastische Abhängigkeiten auf (siehe Abbildung 13). Dieses Ergebnis bestätigt die für dieses Testinstrument angenommene Testletstruktur und untermauert ebenso das hier beschriebene Vorgehen der Bildung von Summenscores zur Punktvergabe für die einzelnen Vignetten. Auf Analyseebene der Vignetten-

summenscores zeigt sich hingegen keine lokale stochastische Abhängigkeit (siehe Abbildung 14). Daher ist für diesen Vignettentest davon auszugehen, dass – unter Kontrolle der Fähigkeitsausprägung der Probanden – die Beantwortung der einzelnen Vignetten unabhängig voneinander erfolgt. Da für diesen Test lokale stochastische Unabhängigkeit vorliegt, entspricht das Produkt der Einzellösungswahrscheinlichkeiten der Vignetten der Wahrscheinlichkeit alle Vignetten korrekt zu beantworten (vgl. Bühner, 2011).

Die graphische Darstellung der Itemkategorieschwellen zeigt, dass diese in den oberen Punktkategorien zum Teil vertauscht sind (siehe Abbildung 15). Einige Autoren halten vertauschte Itemkategorieschwellen bei Partial Credit Modellen hinsichtlich der Modellinterpretation für problematisch (Bühner, 2011; Muraki, 1993). Zum einen wird argumentiert, dass bei vertauschten Itemschwellen nicht (eindeutig) von ordinalen Daten ausgegangen werden kann, da diese nicht in der vorhergesagten Reihenfolge auftreten. Sind zwei Itemschwellen vertauscht – beispielsweise Schwelle 3 und Schwelle 4 – bedeutet das inhaltlich streng genommen, dass es eine geringere Fähigkeit zum Überwinden der vierten Schwelle (und folglich zum Erreichen der vollen Testpunktzahl) erfordert als zum Überwinden der dritten Schwelle. Dies erscheint aus interpretatorischer Sicht jedoch wenig sinnvoll, da zum Erreichen einer höheren Testpunktzahl innerhalb einer Testaufgabe auch eine höhere Fähigkeitsausprägung zugrunde liegen sollte. Zum anderen wird aufgezeigt, dass vertauschte Itemschwellen auf eine unnötig hohe Anzahl an Kategorieabstufungen hinweisen könnten; so wird vermutet, dass die Kategorien zu schmal – und daher wenig trennscharf – seien und dass aus einer Reduktion der Kategorieabstufungen eindeutige Itemschwellen resultieren könnten.

Andere Autoren argumentieren hingegen, dass vertauschte Itemschwellen nicht zwangsläufig ein Problem für die Modellinterpretation darstellen. Auch empirische Gründe können zu vertauschten Schwellen führen, beispielsweise selten besetzte Itemkategorien (Adams, Wu & Wilson, 2012; Masters & Wright, 1997). Dieses Phänomen tritt in der vorliegenden Studie teilweise auf (siehe Tabelle 12).

Adams et al. (2012) weisen darauf hin, dass bei vertauschten Itemschwellen mögliche Ursachen gründlich analysiert werden sollten: In der Wright Map (siehe Abbildung 16) kann abgelesen werden, dass sich die Abstände der Thurstonian Thresholds zwischen den Vignetten teilweise erheblich unterscheiden. Wenn die

Abstände der Thurstonian Thresholds innerhalb einer Vignette sehr gering sind, dann bedeutet das, dass die Studierenden nur eine minimal höhere Fähigkeit benötigen, um die nächsthöhere Itemkategorie zu erreichen. In solchen Fällen verschwimmen die Grenzen zwischen den aufeinanderfolgenden Itemkategorien, so dass die Stufen nicht mehr eindeutig voneinander getrennt abgebildet werden können. Es resultieren u. U. vertauschte Itemschwellen.

An dieser Stelle könnte darüber nachgedacht werden, ob 5 Kategorien (also maximal 4 zu erreichende Punkte) pro Vignette eventuell eine zu feingliedrige Abstufung des zugrunde liegenden Konstrukts – des lesedidaktischen Wissens – darstellen. Es könnte sein, dass die Stufen eindeutiger voneinander trennbar wären, wenn man die oberen Kategorien zusammenfassen und beispielsweise nur maximal 3 Punkte pro Vignette vergeben würde. Dieses Vorgehen würde allerdings einen weiteren Informationsverlust bedeuten; die Vergabe von 0 bis 4 Testpunkten pro Vignette stellt bereits eine Kategorisierung – und damit auch eine Informationsreduktion – der ursprünglich 10 zu bildenden Relationen pro Vignette dar. Daneben wurde der Vignettentest bisher noch nicht bei erfahrenen Lehrkräften eingesetzt. Möglicherweise ergibt sich eine eindeutigere Abgrenzbarkeit in den oberen Itemkategorien erst dadurch, dass auch Probanden mit hoher und transsituational gefestigter Expertise befragt werden (siehe Kapitel 5.5).

Neben einer empirischen Ursache für die vertauschten Itemschwellen kommt ferner eine inhaltliche Ursache in Betracht: Möglicherweise bewerten die Studierenden die Unterrichtssituationen qualitativ anders als die Experten, die zur Normgewinnung befragt wurden. Dies könnte ebenfalls vertauschte Itemschwellen bedingen. Da es in der Forschung zum Expertenwissen gesicherte empirische Evidenzen dafür gibt, dass sich das Wissen von Experten und Novizen qualitativ voneinander unterscheidet (Überblick: Krauss & Bruckmeier, 2014), könnte dieser Überlegung in weiterführenden Studien nachgegangen werden. Hierzu sind allerdings längsschnittliche Untersuchungen notwendig, um die Messinvarianz des Vignettentests für unterschiedliche Expertisegruppen zu untersuchen (Widaman & Reise, 1997).

Die EAP-Reliabilität weist mit 0.59 einen mittelmäßigen Wert auf. Auch andere Autoren berichten bezüglich Skalen zur Erfassung fachdidaktischer Wissensinhalte im Deutschunterricht vergleichbare Reliabilitätsmaße: Pissarek und Schilcher

(2015) berichten für die Aufgabengruppe *Fachdidaktik* ein Cronbachs α von 0.67 (S. 336), Bremerich-Vos und Dämmer (2013) geben für die Skala zur *Literaturdidaktik* eine EAP-Reliabilität von 0.62 (S. 59) an. Da es sich bei dem Vignettentest um ein Testinstrument mit sehr heterogenen Testaufgaben handelt, war ein bedeutsam höherer Wert für das Maß der internen Konsistenz allerdings auch nicht zu erwarten (Bühner, 2011). Für heterogene Testformate erscheint die Überprüfung der Teststabilität als zuverlässigste Methode, um zu einer verlässlichen Schätzung der Reliabilität zu kommen³⁹ (vgl. Lienert & Raatz, 1998, S. 202). Mithilfe der Spearman-Brown Formel wurde festgestellt, dass die Reliabilität des Vignettentests bei .79 liegen würde, wenn alle der ursprünglich erstellten 31 Testaufgaben eingesetzt worden wären (unter der Annahme, dass die zusätzlichen Testaufgaben ähnliche psychometrische Eigenschaften aufweisen wie die eingesetzten Testaufgaben) (Moosbrugger & Kelava, 2012). Von einer deutlichen Erweiterung des Vignettentests durch die Aufnahme vieler zusätzlicher Vignetten wird allerdings abgesehen, da sich die Bearbeitungsdauer deutlich verlängern würde, was einem Hauptanliegen dieser Studie – der Ökonomie bei der Testdurchführung – nicht zuträglich wäre.

12.4 Validität

Die inhaltliche Validität des Testverfahrens wurde durch einen mehrstufigen Prozess mit qualitativen und quantitativen Expertenbefragungen sichergestellt. Um weitere Aspekte der Validität des Vignettentests zu untersuchen, wurde ein Kontrollgruppenvergleich durchgeführt. Des Weiteren wurde überprüft, ob Unterschiede im Ausmaß des lesedidaktischen Wissens in Abhängigkeit von der Studiendauer bei den Lehramtsstudierenden zu verzeichnen sind. Ferner wurde untersucht, inwiefern die praktischen Erfahrungen, die die Studierenden bereits erworben haben, die Leistung im Testverfahren vorhersagen können.

12.4.1 Kontrollgruppenvergleich

Der Kontrollgruppenvergleich liefert eingeschränkte Hinweise auf die Validität des Vignettentests: Die Deutschstudierenden schneiden im Testverfahren erwar-

³⁹ Unter 15 werden die geplanten weiterführenden Analysen zur Untersuchung der Reliabilität des Vignettentests dargestellt.

tungsgemäß am besten ab. Die einfachen geplanten Kontraste zu den Psychologiestudierenden und den Germanistikstudierenden erreichen die statistische Signifikanz. Allerdings schneiden die Deutschstudierenden im Testverfahren nur marginal signifikant besser ab als die Lehramtsstudierenden fachfremder Lehrämter. Auch sind die Mittelwertunterschiede zwischen den verglichenen Gruppen als eher gering zu bezeichnen (durchschnittlich 22.7 Testpunkte bei den Germanistikstudierenden und durchschnittlich 25.5 Testpunkte bei den Deutschstudierenden).

Möglicherweise fragt der Vignettentest zu allgemeine (kognitive oder didaktische) Wissensinhalte ab, die nicht spezifisch genug für die Lesedidaktik sind. Allerdings wurden bereits Indikatoren der allgemeinen kognitiven Leistungsfähigkeit genutzt, um diese Effekte zu kontrollieren. Des Weiteren haben alle Vignetten ein mehrstufiges Expertenrating durchlaufen, sodass die inhaltliche Validität des Testverfahrens als gesichert angesehen werden kann. Die fachdidaktische Relevanz wurde bei allen zwölf Vignetten als überdurchschnittlich hoch bewertet (siehe Tabelle 12).

Da die Förderung von Lesekompetenz ein Anliegen aller Schulfächer in der Sekundarstufe I darstellt (Artelt & Dörfler, 2010) ist es auch denkbar, dass Lehramtsstudierende ohne das Fach Deutsch während ihres Studiums mit dem Themenbereich „Lesedidaktik“ in Berührung kommen. Dies ist insbesondere für Lehramtsstudierende mit Fremdsprachen zu erwarten. Es berichten jedoch nur 3 % der Lehramtsstudierenden ohne das Fach Deutsch, dass sie bereits Lehrerfahrung im Leseunterricht erworben und 5.8 %, dass sie bereits ein Seminar zur Lesedidaktik besucht haben. Allerdings wurde im Fragebogen nicht der Besuch von (lese)didaktischen Seminaren zu Fremdsprachen erfasst. D. h. dass möglicherweise mehr als 5.8 % der Studierenden ohne das Fach Deutsch ein Seminar zur Lesedidaktik besucht haben, das sich jedoch auf den Fremdsprachenunterricht bezogen hat. Der Besuch lesedidaktischer Seminare in Fremdsprachen könnte sich womöglich positiv auf die Testleistung auswirken. In nachfolgenden Studien sollte daher explizit abgefragt werden, ob die Studierenden (lese)didaktische Seminare zu Fremdsprachen besucht haben.

Daneben ist anzumerken, dass sich die Sub-Stichproben, die für diesen Kontrollgruppenvergleich herangezogen wurden, hinsichtlich Stichprobenumfang und Repräsentativität (z. B. bezüglich der Studierenden in verschiedenen Fachsemestern)

deutlich unterscheiden (siehe Tabelle 11). Für anschließende Analysen sollte daher versucht werden, die Kontrollgruppen durch nachträgliche Erhebungen zu erweitern.

Fraglich ist nun, ob das Ergebnis des Kontrollgruppenvergleichs als eine Einschränkung der diskriminanten Validität des Vignettentest aufgefasst werden sollte. In der Literatur werden bislang keine Kontrollgruppenvergleiche berichtet, die sich auf fachdidaktische Wissensinhalte von Lehramtsstudierenden (also auf Novizen) aus verschiedenen Studiengängen beziehen. Daher liegen keine Referenzmaße vor, sodass ein entsprechender Vergleich zu ähnlichen Studien (momentan) nicht möglich ist.

Es wurde allerdings anhand von Expertenbefragungen untersucht, inwiefern Shulmans (1987) Annahme, dass PCK *der* Wissensinhalt ist, der Lehrkräfte von Fachwissenschaftlern und pädagogischen Fachkräften unterscheidet. Hierzu wurden in verschiedenen Domänen Befragungen mit erfahrenen Lehrkräften, Fachwissenschaftlern und Pädagogen durchgeführt. In den Naturwissenschaften liegen gesicherte Befunde darüber vor, dass Lehrkräfte in Testverfahren zu PCK besser abschneiden als Fachwissenschaftler und Pädagogen (Jütter & Neuhaus, 2013), allerdings auch dort mit teilweise geringen Mittelwertunterschieden (Witner & Tepner, 2011). Wendet man diese Ergebnisse auf die vorliegende Studie an, kann festgestellt werden, dass sich Deutschlehramtsstudierende (vergleichbar zu „Lehrkräften“) hinsichtlich des lesedidaktischen Wissens signifikant von den Psychologiestudierenden (vergleichbar zu „Pädagogen“) und den Germanistikstudierenden (vergleichbar zu „Fachwissenschaftlern“) unterscheiden. Damit sind diese Ergebnisse – selbstverständlich in eingeschränkter Art und Weise – anschlussfähig an empirische Evidenzen aus vorherigen Studien.

Ob der nicht signifikante Unterschied zwischen Deutschstudierenden und den Studierenden fachfremder Lehrämter bezüglich des lesedidaktischen Wissens als Indikator für eine mangelnde diskriminante Validität des Vignettentests gelten muss, kann erst dann abschließend beurteilt werden, wenn nachfolgende Studien Ergebnisse zur Spezifität professioneller Wissensinhalte bei Lehramtsstudierenden verschiedener Domänen vorgelegt haben.

12.4.2 Anstieg des lesedidaktischen Wissens über das Studium (im Querschnitt)

Es wurde erwartet, dass das lesedidaktische Wissen von Deutschstudierenden über das Studium hinweg (querschnittlich) ansteigt. In der vorliegenden Studie wurde diese Annahme trendanalytisch überprüft.

In einem latenten Regressionsmodell zeigte sich, dass die Studiendauer (operationalisiert über die Semesteranzahl) das lesedidaktische Wissen der Studierenden vorhersagen kann. Dieses Ergebnis spricht für die curriculare Validität, da die Deutschlehramtsstudierenden über ihr Studium hinweg zunehmend lesedidaktisches Wissen erwerben (sollten). Allerdings weist die hohe Residualvarianz des latenten Faktors „lesedidaktisches Wissen“ darauf hin, dass es neben der Semesteranzahl weitere einflussreiche Faktoren auf das lesedidaktische Wissen gibt (siehe auch Kapitel 12.4.3).

In einem weiterführenden Analyseschritt wurde für die verschiedenen Substichproben (Primarstufenstudierende, Sekundarstufenstudierende, Sonderpädagogikstudierende, Gymnasiallehramtsstudierende, Studierende fachfremder Lehrämter, Psychologiestudierende, Germanistikstudierende) überprüft, inwiefern eine Veränderung des lesedidaktischen Wissens von Beginn (Semester 1 – 3) bis zum Ende des Studiums (Semester 4 – 8) festzustellen ist.

Für alle Deutschlehramtsstudiengänge war ein deskriptiver Anstieg des lesedidaktischen Wissens vom Beginn bis zum Ende des Studiums zu beobachten. Diese Ergebnisse sind analog zu anderen Studien, die fachdidaktische Wissensinhalte von Deutschstudierenden (Bremerich-Vos & Dämmer, 2013) oder fachdidaktische Wissensinhalte in anderen Domänen (Schmelzing, Wüsten, Sandmann & Neuhäus, 2010) querschnittlich untersucht haben. Jedoch erreicht dieser Anstieg nur bei den Sekundarstufenlehramtsstudierenden die statistische Bedeutsamkeit. Dieses Ergebnis kann dahingehend interpretiert werden, dass der Vignettentest insbesondere professionelle fachdidaktische Wissensinhalte abfragt, die im Rahmen eines Deutschlehramtsstudiums für die Sekundarstufe I erworben werden; dies liegt nahe, da der Test ausschließlich Vignetten umfasst, die sich auf Unterrichtssituationen in der Sekundarstufe I beziehen.

Für die Kontrollgruppen ist keinerlei Zuwachs des lesedidaktischen Wissens zu beobachten. Letzteres spricht für die Validität des Vignettentests: Es ist plausibel anzunehmen, dass innerhalb der Deutschlehramtsstudiengänge ein Zuwachs an lesedidaktischem Wissen stattfindet, wohingegen dies in den Kontrollgruppen nicht der Fall sein sollte.

Die Sekundarstufenlehramtsstudierenden wurden näher untersucht: Hier ist ein kleinschrittiger aber kontinuierlicher Anstieg des lesedidaktischen Wissens vom ersten Semester bis zum Referendariat (im Querschnitt) zu beobachten. Der beobachtete „Einbruch“ des Testscores im siebten Semester ist zu vernachlässigen, da hier aufgrund der Ziehung einer Gelegenheitsstichprobe in dieser Gruppe zufällig nur vier Studierende befragt wurden (siehe Tabelle 11⁴⁰).

Für den Vergleich der Testpunktzahlen der Deutschlehramtsstudierenden im ersten Semester (22.44 Testpunkte) und der Referendare (26.86 Testpunkte) ergibt sich ein starker Effekt von $d = 0.81$ (Cohen, 1988). Trotz des starken statistischen Effekts ist ein Anstieg von 4.5 Testpunkten als gering bis mittelmäßig zu bezeichnen. Für diesen relativ geringen Anstieg kann es unterschiedliche Gründe geben: Es könnte sein, dass für Deutschlehramtsstudierende tatsächlich ein vergleichsweise geringer Anstieg des lesedidaktischen Wissens über das Studium hinweg zu verzeichnen ist. PCK ist – wie unter 5.3 beschrieben – als professioneller Wissensinhalt mit teilweise impliziten Anteilen zu verstehen, dessen Aufbau u. a. als erfahrungsabhängig gilt. Da Lehramtsstudierende während des Hochschulstudiums kaum eigene Lehrerfahrung erwerben, ist ein deutlicher Anstieg an PCK daher möglicherweise auch nicht zu erwarten. Eine anschließende Fragestellung an die hier präsentierte Studie wäre den Verlauf von CK bezüglich Lesedidaktik und Leseprozess über das Studium hinweg zu untersuchen (querschnittlich und längsschnittlich); da es sich bei CK um ausschließlich explizites Wissen handelt, werden hier stärker Zuwächse erwartet als für den Anstieg von PCK.

Einschränkend ist hier jedoch darauf hinzuweisen, dass es sich bei der vorliegenden Stichprobe um eine Querschnittserhebung handelt und nicht um eine echte längsschnittliche Untersuchung der Kompetenzentwicklung der Studierenden.

⁴⁰ Es wurden 8 Studierende befragt, die sich in Fachsemester 5 befinden. Ansonsten ist für alle Semester $N > 10$.

Bemerkenswert ist, dass sich gerade für die Sekundarstufenstudierenden der deutlichste Anstieg des lesedidaktischen Wissens empirisch abbilden lässt: Die Testaufgaben beziehen sich ja explizit auf Unterrichtssituationen aus der Sekundarstufe. Offensichtlich ist der Vignettentest insbesondere für die Zielstichprobe der Sekundarstufenstudierenden sensitiv. Inwiefern sich dieser Effekt – besondere Sensitivität von Sekundarstufenvignetten bei Sekundarstufenstudierenden – replizieren lässt, muss weitergehend untersucht werden.

12.4.3 Bedeutung der praktischen Erfahrungen für die Entwicklung des lesedidaktischen Wissens

In der Literatur wird die Annahme vertreten, dass spezifische Lerngelegenheiten im Lehramtsstudium (König & Seifert, 2012) sowie die Lehrerfahrung (van Driel, Beijaard & Verloop, 2001), die u. a. im Rahmen von studienbegleitenden Praktika erworben wird, bedeutsam für die Entwicklung von fachdidaktischen Wissensinhalten von angehenden Lehrkräften sind. Neben der Untersuchung, inwiefern das lesedidaktische Wissen der Studierenden durch den unspezifischen Einflussfaktor Studiendauer vorhergesagt wird, war daher eine weitere interessierende Fragestellung, welche Rolle die praktischen Erfahrungen der Studierenden bei der Entwicklung des lesedidaktischen Wissens spielen. Unter praktischen Erfahrungen wurden in der vorliegenden Studie die eigene Lehrerfahrung verstanden, die die Studierenden bereits erworben haben sowie der Besuch spezifischer Seminare, die konkrete lesedidaktische Vorgehensweisen thematisieren.

Zunächst wurde untersucht, ob bereits erworbene Lehrerfahrung der Studierenden das lesedidaktische Wissen vorhersagen kann. Die Ergebnisse der latenten Regression zeigen hier keine statistische Bedeutsamkeit des Prädiktors Lehrerfahrung. Spezifische Lehrerfahrungen hingegen, also Lehrerfahrung *im Leseunterricht*, können die Testleistung signifikant vorhersagen. Dieses Ergebnis spricht deutlich für die Validität des Testverfahrens und für die Annahme, dass durch diesen Test auch implizite Anteile des fachdidaktischen Wissens im Leseunterricht erfasst werden, die für das Unterrichtshandeln bedeutsam sind: Nur wenn die Studierenden tatsächlich Lehrerfahrung im Leseunterricht gesammelt haben, schlägt sich diese Lehrerfahrung in der Testleistung nieder. Lehrerfahrung im Allgemei-

nen, beispielsweise in den weiteren Fächern, beeinflusst das lesedidaktische Wissen der Studierenden nicht.

Weiterhin konnte gezeigt werden, dass der Besuch von spezifischen lesedidaktischen Seminaren die Testleistung signifikant vorhersagen kann. Auch dieses Ergebnis liefert Hinweise auf die Validität des Vignettentests: Es werden hier offensichtlich lesedidaktische Wissensinhalte abgefragt, die im Rahmen spezifischer Veranstaltungen an der Hochschule vermittelt werden.

Da die praktischen Erfahrungen – Lehrerfahrung und der Besuch spezifischer lesedidaktischer Seminare – offensichtlich einen Einfluss auf die Testleistung der Studierenden und somit auf die Entwicklung des lesedidaktischen Wissens haben, sollte in einem weiteren Analyseschritt untersucht werden, inwiefern die praktischen Erfahrungen neben der bloßen Studiendauer (operationalisiert über die Semesteranzahl) die Testleistung der Studierenden vorhersagen können. Hierzu wurde ein latentes Regressionsmodell berechnet, das eine sehr gute Modellpassung aufweist (Hu & Bentler, 1998). Hierbei zeigt sich ein Mediationseffekt: Die Studiendauer kann zwar das Ausmaß an praktischen Erfahrungen vorhersagen (je länger das Studium dauert, desto eher erwerben die Studierenden Lehrerfahrung und besuchen lesedidaktische Seminare), jedoch nicht das lesedidaktische Wissen der Studierenden. Die praktischen Erfahrungen sind jedoch ein statistisch bedeutsamer Prädiktor für das lesedidaktische Wissen. Folglich lässt dieses Ergebnis darauf schließen, dass nicht die bloße Studiendauer einen Einfluss auf die Entwicklung des lesedidaktischen Wissens nimmt, sondern die praktischen Erfahrungen der Studierenden (Abd-El-Khalick, 2006; Hascher, 2006; Kunina-Habenicht et al., 2013; van Driel et al., 2001; van Driel & Berry, 2012).

Die Residualvarianz des latenten Faktors „lesedidaktisches Wissen“ weist mit 0.93 einen hohen Wert auf. Das bedeutet, dass neben den praktischen Erfahrungen noch viele weitere Einflussgrößen auf den latenten Faktor wirken. Dieses Ergebnis war zu erwarten, da die Entwicklung von PCK als ein äußerst komplexer Integrationsprozess verschiedener professioneller Wissensinhalte zu verstehen ist (Henze, van Driel & Verloop, 2008; van Driel & Berry, 2012). Unter 5.5.2 sind die verschiedenen Einflussfaktoren auf die Entwicklung von PCK dargestellt: So gelten neben der praktischen Lehrerfahrung der Erwerb von CK, Erfahrungen aus der eigenen Schulzeit sowie die Zusammenarbeit mit Kollegen als bedeutsam für

den Aufbau von PCK. Inwiefern diese verschiedenen Einflussfaktoren zusammenwirken und sich möglicherweise gegenseitig bedingen ist bislang kaum untersucht worden und sollte Gegenstand weiterführender Untersuchungen sein.

Die berichteten Ergebnisse können ferner in Bezug zu den Überlegungen von Neuweg (2014) diskutiert werden (siehe Kapitel 5.4): Neuweg nimmt an, dass Wissensbereich 2 – in der vorliegenden Studie das lesedidaktische Wissen – sowohl durch Wissensbereich 1 (explizites Ausbildungswissen) als auch durch Lernen durch Erfahrung beeinflusst wird. In dieser Studie wurde Wissensbereich 1 über die Studiendauer operationalisiert; Lernen durch Erfahrung wurde über die bereits erworbene Lehrerfahrung im Leseunterricht sowie den Besuch lesedidaktischer Seminare operationalisiert. Es zeigt sich, dass nur letztere einen statistisch bedeutsamen Einfluss auf das lesedidaktische Wissen haben. In Neuwegs Terminologie gesprochen konnte in dieser Studie ausschließlich ein Einfluss von Lernen durch Erfahrung auf Wissensbereich 2 festgestellt werden, jedoch nicht von Wissensbereich 1 auf Wissensbereich 2. Dieses Ergebnis hängt möglicherweise mit der z. T. eingeschränkten Operationalisierung der Konstrukte aus Neuwegs Modell zusammen. Wissensbereich 1 wurde hier ausschließlich über die Semesteranzahl erfasst. Diese lässt nur bedingt Rückschlüsse auf das Ausmaß an Ausbildungswissen zu, das die Studierenden im Laufe ihrer Ausbildung erwerben. Eine differenziertere Erfassung von Wissensbereich 1 wäre für folgende Studien daher wünschenswert, beispielsweise durch die Erhebung von CK.

12.5 Zusammenhänge des lesedidaktischen Wissens zu verwandten Konstrukten

Die professionelle Kompetenz von Lehrkräften wird als Bündel bestehend aus verschiedenen kognitiven sowie affektiven Teilkompetenzen konzeptualisiert (Kunter, Baumert et al., 2011). Eine weitere Fragestellung adressiert daher die Untersuchung von (Wirk)Zusammenhängen zwischen den einzelnen Kompetenzaspekten.

In der vorliegenden Studie wurden Zusammenhänge zwischen dem lesedidaktischen Wissen und kognitiven sowie affektiven Variablen, die sich in das Modell der professionellen Kompetenz von Lehrkräften einordnen lassen, z. T. explorativ untersucht. So sollte eine Einbettung des lesedidaktischen Wissens in den breite-

ren Kontext der professionellen Kompetenz von Lehrkräften ermöglicht werden. Des Weiteren wurden Zusammenhänge zwischen dem lesedidaktischen Wissen und den BIG-5 überprüft.

Der ausführlichen Diskussion der Zusammenhänge zwischen dem lesedidaktischen Wissen der Studierenden und den erhobenen kognitiven und affektiven Konstrukten ist vorwegzunehmen, dass durchgängig geringe korrelative Zusammenhänge gefunden wurden (siehe Tabelle 13). Allerdings kann davon ausgegangen werden, dass es sich hierbei um zuverlässige Werte handelt, da Ergebnisse aus Monte-Carlo-Simulationsstudien zeigen, dass sich stabile Schätzungen für bivariate Korrelationen ab $N = 250$ ergeben (Schönbrodt & Perugini, 2013). Die Reliabilitätsschätzungen der erhobenen Konstrukte (gemessen an Cronbachs α , siehe Kapitel 10.3) sind abgesehen von den Skalen zu den BIG-5 als gut bis sehr gut zu bezeichnen.

Zunächst wurden Zusammenhänge zwischen den kognitiven Variablen und dem lesedidaktischen Wissen untersucht: Hier zeigen sich korrelative Zusammenhänge zwischen dem lesedidaktischen Wissen und den Indikatoren für allgemeine kognitive Leistungen (hier operationalisiert über die Abiturdurchschnittsnote und sprachliches Denken). Auch andere Autoren berichten Zusammenhänge zwischen der Abiturdurchschnittsnote und dem Abschneiden in fachbezogenen Wissenstests bei Lehramtsstudierenden (Blömeke, Felbrich et al., 2008; König & Seifert, 2012).

Aufgrund dieses Ergebnisses wurden alle weiteren Analysen zu Zusammenhängen zwischen dem lesedidaktischen Wissen und den verwandten Konstrukten als partielle Korrelationen mit den Kovariaten „Abiturdurchschnitt“ und „sprachliches Denken“ durchgeführt (vgl. Kunina-Habenicht et al., 2013). So sollten Scheinkorrelationen zwischen den Konstrukten verhindert werden, die sich möglicherweise durch den Einfluss der allgemeinen kognitiven Leistungsfähigkeit auf die beiden zu korrelierenden Konstrukte ergeben könnten.

Zuerst wurde der Zusammenhang zwischen dem lesedidaktischen Wissen und PK untersucht. Hier zeigt sich hypothesenkonform ein positiver Zusammenhang. Dieser fiel mit $r = 0.15$ etwas geringer aus, als vergleichbare Studien aus den Naturwissenschaften berichten (Kirschner, 2013; Lin et al., 2013; Paulick et al., 2016).

Nichtsdestoweniger kann hier anschließend an entsprechende Vorarbeiten ein Zusammenhang zwischen den beiden professionellen Wissensdomänen PCK und PK gezeigt werden, der sich auch nach der Kontrolle der allgemeinen kognitiven Leistungsfähigkeit zeigt. Daher ist davon auszugehen, dass die beiden professionellen Wissensinhalte tatsächlich einen Zusammenhang aufweisen.

Bezüglich der Zusammenhänge des lesedidaktischen Wissens zu affektiven Variablen ergab sich ein differenziertes Bild: Die Skalen zu lesedidaktischen Überzeugungen weisen allesamt statistisch bedeutsame Zusammenhänge mit dem lesedidaktischen Wissen auf. Eine konstruktivistische lesedidaktische Überzeugung hängt – wie erwartet – positiv mit dem Testscore zusammen, eine direkt-transmissive Überzeugung negativ. Studierende, die dem lesedidaktischen Unterricht gegenüber eine konstruktivistische Überzeugung einnehmen, beantworten die Testaufgaben daher eher im Sinne der Experten als Studierende mit einer direkt-transmissiven Überzeugung. Offensichtlich geht eine konstruktivistische Sichtweise der Studierenden bezüglich der Vermittlung von Lesestrategien mit einem höheren Ausmaß an lesedidaktischem Wissen einher. Eine konstruktivistische Sichtweise auf Lehr-Lernprozesse sieht die Schüler als aktiv am Lernprozess beteiligt an und nicht als passive Rezipienten von fachlichen Inhalten; eine konstruktivistische Ansicht bezüglich Lehr-Lernprozessen liegt auch verschiedenen deutschdidaktische Vorgehensweisen zu Förderung von Lesekompetenz zugrunde: Bei der Förderung von metakognitiven Lesestrategien wird beispielsweise darauf abgezielt, selbstregulative Fähigkeiten der Schüler zur eigenständigen Planung und Überprüfung des eigenen Leseprozesses zu stärken.

In der Literatur liegen Evidenzen vor, dass eine konstruktivistische Überzeugung der Lehrkraft mit besseren Schülerleistungen einhergeht (Behrmann & Souvignier, 2015; Dubberke et al., 2008). Man könnte dieses Ergebnis daher weiterführend so interpretieren, dass durch den positiven Zusammenhang zwischen dem lesedidaktischen Wissen und der konstruktivistischen Überzeugung im Leseunterricht ein erster Hinweis darauf vorliegen könnte, dass das lesedidaktische Wissen – ebenso wie eine konstruktivistische Überzeugung – einen positiven Einfluss auf die Schülerleistung haben könnte. Diese Überlegung bleibt jedoch sehr vage und ist durch anschließende Studien zu überprüfen. An dieser Stelle ist jedoch anzumerken, dass es sich bei der hier beschriebenen Stichprobe um Deutschlehramts-

studierende handelt, d. h. dass deren Überzeugungen bezüglich der Vermittlung von Lesestrategien vermutlich noch nicht gefestigt vorliegen wie beispielsweise bei erfahrenen Lehrkräften. Des Weiteren ist damit zu rechnen, dass konstruktivistische bzw. direkt-transmissive Überzeugungen in der Berufseinstiegsphase noch veränderbar sind (Behrmann & Souvignier, 2013).

Bezüglich des berufsbezogenen Selbstkonzeptes der Studierenden wurden in der vorliegenden Studie die drei Skalen „Selbstkonzept Fach“, „Selbstkonzept Diagnostik“ und „Selbstkonzept Erziehung“ abgefragt. Es konnten keine Zusammenhänge zwischen dem lesedidaktischen Wissen und den Skalen zum berufsbezogenen Selbstkonzept festgestellt werden.

Dieses Ergebnis ist überraschend, da in einer vorhergehenden Arbeit korrelative Zusammenhänge zwischen den Noten von Lehramtsstudierenden (also Studienleistungen, die sich auf professionelle Wissensinhalte beziehen) im ersten und zweiten Unterrichtsfach mit den Skalen „Selbstkonzept Fach“ und „Selbstkonzept Erziehung“ gezeigt werden konnten (Retelsdorf et al., 2014b). Möglicherweise bleiben die erwarteten Zusammenhänge an dieser Stelle aus, da es sich bei den Noten in den Unterrichtsfächern eher um Indikatoren für CK handelt; in dieser Studie wurde mit PCK eine andere Domäne des professionellen Lehrerwissens erfasst.

Passend zu dieser Überlegung berichten Paulick et al. (2016) eine Untersuchung zu Zusammenhängen zwischen dem berufsbezogenen Selbstkonzept angehender Biologie- und Physiklehrkräfte und den Wissensinhalten CK, PCK und PK. Es wurde hier untersucht, inwiefern es spezifische Selbstkonzepte für unterschiedliche professionelle Wissensbereiche gibt. Es wurden daher Skalen zur Erfassung des berufsbezogenen Selbstkonzeptes adaptiert und eingesetzt, die sich spezifisch auf die einzelnen Domänen des professionellen Wissens von Lehrkräften beziehen: Selbstkonzept bezüglich CK, Selbstkonzept bezüglich PCK und Selbstkonzept bezüglich PK. Diese drei Skalen lassen sich in einem Modell mit drei Faktoren empirisch abbilden, was darauf hinweist, dass von spezifischen Selbstkonzeptbereichen bezüglich dieser professionellen Wissensinhalte auszugehen ist. Tatsächlich weist das Selbstkonzept bezüglich PCK in dieser Studie einen höheren korrelativen Zusammenhang zu PCK auf (angehende Biologielehrkräfte: $r = 0.26$, $p \leq 0.001$, angehende Physiklehrkräfte: $r = 0.30$, $p \leq 0.001$) als zu CK (angehende Biologielehrkräfte: $r = 0.03$, n. s., angehende Physiklehrkräfte:

$r = 0.05$, n. s.) und PK (angehende Biologielehrkräfte: $r = 0.13$, $p \leq 0.001$, angehende Physiklehrkräfte: $r = 0.21$, $p \leq 0.001$).

Möglicherweise blieben die Zusammenhänge zwischen dem lesedidaktischen Wissen und den Skalen zum berufsbezogenen Selbstkonzept aus, da diese zu unspezifisch für das lesedidaktische Wissen sind. Es hätte sich womöglich ein Zusammenhang zwischen dem lesedidaktischen Wissen und dem berufsbezogenen Selbstkonzept der Studierenden ergeben, wenn letzteres spezifischer erfasst worden wäre, beispielsweise in Form einer Skala zum Selbstkonzept bezüglich des fachdidaktischen Wissens im Leseunterricht.

In vorherigen Studien ergaben sich zwar Hinweise auf Zusammenhänge zwischen CK und Aspekten des berufsbezogenen Selbstkonzeptes; bezüglich PCK konnten solche Zusammenhänge in dieser Studie nicht gezeigt werden. Dieses Ergebnis betont daher nochmals die Spezifität der beiden professionellen Wissensdomänen CK und PCK (Baumert et al., 2010).

Bezüglich der verschiedenen Skalen zur Berufswahlmotivation der Studierenden wurde erwartet, dass Aspekte der intrinsischen Motivation einen positiven Zusammenhang mit dem lesedidaktischen Wissen aufweisen sollten, extrinsische Aspekte einen negativen Zusammenhang. Hier wies nur die Skala Pädagogisches Interesse – in erwarteter Richtung – einen statistisch bedeutsamen Zusammenhang mit dem lesedidaktischen Wissen auf. Warum die anderen Skalen zur Berufswahlmotivation keine Zusammenhänge mit dem lesedidaktischen Wissen aufweisen, kann nur spekuliert werden. Naheliegend ist die Vermutung, dass Aspekte der Berufswahlmotivation tatsächlich nur in geringem bzw. keinem Zusammenhang mit dem lesedidaktischen Wissen stehen. Es könnte auch sein, dass an dieser Stelle Skalen zu anderen motivationalen Aspekten gewinnbringender eingesetzt werden könnten, beispielsweise eine Skala zu *teacher enthusiasm*. Dies ist möglicherweise ein motivationales Konstrukt, das mit dem Zielkonstrukt „lesedidaktisches Wissen“ in höherem Ausmaß in Zusammenhang steht, da es sich bei *teacher enthusiasm* sowie dem lesedidaktischen Wissen um Konstrukte handelt, die sich auf die Situation im Unterricht beziehen. Die Berufswahlmotivation ist hingegen als ein motivationales Konstrukt anzusehen, das der Entwicklung der professionellen Kompetenzen und Wissensinhalten von angehenden Lehrkräften vorangestellt ist.

Befunde aus anderen Studien weisen darauf hin, dass intrinsische Motivationsaspekte zu Beginn des Studiums die Anstrengung und Beharrlichkeit vorhersagen, mit der die Studierenden sich ihrem Studium widmen (Watt & Richardson, 2007). Es ist allerdings unklar, inwiefern Anstrengung und Beharrlichkeit im Studium auch zu einem Erwerb an (implizitem) lesedidaktischem Wissen führen, da dieses ja nicht wie fachwissenschaftliche Inhalte „auswendig gelernt“ werden kann, sondern durch zunehmende eigene Erfahrungen entwickelt und gefestigt wird (Lee et al., 2007). Konkretere Befunde, die Zusammenhänge zwischen Aspekten der Berufswahlmotivation und der Studienleistung der Lehramtsstudierenden bzw. die Entwicklung professioneller Wissensinhalte untersuchen, liegen bisher nicht vor. An dieser Stelle sind weitere Forschungsarbeiten notwendig, um gesicherte Aussagen über die prädiktive Bedeutung von Berufswahlmotiven treffen zu können (Kunter, 2014).

Die Untersuchung von Zusammenhängen zwischen weiteren, globalen Persönlichkeitseigenschaften und dem lesedidaktischen Wissen zeigen unerwarteterweise einen statistisch bedeutsamen Zusammenhang zwischen dem lesedidaktischen Wissen und der Persönlichkeitseigenschaft Verträglichkeit. Eigentlich waren Zusammenhänge des lesedidaktischen Wissens mit den Eigenschaften Extraversion und Gewissenhaftigkeit vermutet worden, da diese einen Zusammenhang von großer praktischer Bedeutsamkeit zur pädagogischen Handlungskompetenz im Praktikum aufweisen (Mayr, 2014); die vermuteten Zusammenhänge blieben allerdings aus.

Die Überlegung, dass die ausbleibenden Zusammenhänge zwischen Extraversion und Gewissenhaftigkeit auf die geringe Reliabilität der beiden Skalen zurückzuführen sein könnten (Extraversion: $\alpha = 0.65$, Gewissenhaftigkeit: $\alpha = 0.67$), wird dadurch entkräftet, dass die Skala zu Verträglichkeit eine noch geringere Reliabilität aufweist (Verträglichkeit: $\alpha = 0.55$).

Es liegen Befunde vor, dass die Persönlichkeitseigenschaft Verträglichkeit Zusammenhänge von mittlerer praktischer Bedeutsamkeit zu Lernstrategien im Studium, pädagogischer Handlungskompetenz und Zufriedenheit im Studium aufweist (Mayr, 2014). Diese Zusammenhänge sind jedoch für die Persönlichkeitseigenschaften Extraversion und Gewissenhaftigkeit stärker ausgeprägt, sodass nur

spekuliert werden kann, warum gerade die Skala Verträglichkeit einen statistisch bedeutsamen Zusammenhang mit dem lesedidaktischen Wissen aufweist.

Die Items zur Skala Verträglichkeit sind folgendermaßen formuliert: „Ich bin jemand, der manchmal etwas grob zu anderen ist. (-)“, „Ich bin jemand, der verzeihen kann.“ und „Ich bin jemand, der rücksichtsvoll und freundlich mit anderen umgehen kann.“. Diese Items spiegeln möglicherweise über die Persönlichkeitseigenschaft Verträglichkeit hinaus auch einen wertschätzenden Umgang der Lehramtsstudierenden mit ihren (zukünftigen) Schülern wider. Dieser Überlegung folgend sollte ein wertschätzender Umgang mit Schülern mit einem höheren Ausmaß an lesedidaktischem Wissen einhergehen. Dies würde auch mit dem Ergebnis übereinstimmen, dass der Berufswahlmotivationsaspekt „Pädagogisches Interesse“ ebenfalls einen positiven Zusammenhang mit dem lesedidaktischen Wissen aufweist. Ob dieser Befund so haltbar ist, müssen weitere Studien zeigen.

13 Limitationen der vorliegenden Studie

Bei der Interpretation der Ergebnisse der vorliegenden Studie sind einige Einschränkungen bezüglich der Testentwicklung sowie des Testformates, der psychometrischen Eigenschaften der Testaufgaben und der Durchführung der Studie zu beachten, die in diesem Kapitel vorgestellt werden.

Der Vignettentest soll zwar lesedidaktisches Wissen erfassen, allerdings beziehen sich die zwölf Unterrichtssituationen, aus denen das entwickelte Testinstrument besteht, selbstverständlich nur auf eine Auswahl an möglichen Situationen aus dem Leseunterricht. Die Beurteilung dieser zwölf Unterrichtssituationen wird hier als stellvertretend für den Wissensbereich des lesedidaktischen Wissens angesehen und kann daher nicht alle Bereiche des Leseunterrichts abbilden. Darüber hinaus mussten die Vignetten zu einem Teilbereich der lesedidaktischen Forschung (Bereich „Aufgabenschwierigkeit“) aus dem Testinstrument ausgeschlossen werden, da sich diese Vignetten im Prozess der inhaltlichen Validierung durch die mehrstufige qualitative und quantitative Expertenbefragung nicht bewährt haben (siehe auch Diskussion unter 12.1).

Eine weitere Einschränkung bei der Interpretation der Ergebnisse ist, dass die Unterrichtssituationen in den Vignetten stark vereinfacht und sehr knapp dargestellt werden. Nur so kann die Bearbeitung von zwölf Unterrichtsvignetten im Rahmen einer Testung (in der Regel eine Veranstaltung von 90 Minuten) zeitlich umgesetzt werden. Darüber hinaus wird für die Erstellung von Testaufgaben und Items eine möglichst einfache und knappe Formulierung in Alltagssprache empfohlen, um zu gewährleisten, dass alle Probanden die Testaufgaben gleich verstehen (Thielsch & Brandenburg, 2012). Das tatsächliche Unterrichtsgeschehen ist zweifelsohne um ein Vielfaches komplexer als in den Vignetten dargestellt und es kann nicht davon ausgegangen werden, dass die korrekte Beantwortung einer Unterrichtsvignette auch direkt zu professionellem Handeln im Klassenzimmer führt. Hier sind weitere Studien notwendig, um die prädiktive Validität des Vignettentests auf das Handeln im Unterricht zu untersuchen.

Bezüglich der psychometrischen Eigenschaften der Testaufgaben ist anzumerken, dass alle Vignetten einen ähnlichen allgemeinen Schwierigkeitsindex δ . aufweisen (siehe Abbildung 15); es liegen also weder deutlich schwierige noch leichte Testaufgaben im Testinstrument vor. Dies wäre allerdings von Vorteil für die psychometrischen Eigenschaften des Vignettentests, da schwierige und leichte Items eine Differenzierung in den Randbereichen der Skala ermöglichen (Bühner, 2004). Um eine Differenzierungsfähigkeit zwischen Gruppen unterschiedlicher Leistungsniveaus zu gewährleisten sollten sowohl schwierige als auch leichte Items im Testinstrument präsentiert werden (Moosbrugger & Kelava, 2012). Es könnte sein, dass sich mithilfe von Testaufgaben mit offenem Antwortformat höhere Itemschwierigkeiten ergeben würden; bei offenen Formaten müssen die Probanden mögliche Lösungen nämlich selbst generieren und formulieren, da hier keine zu bewertenden Handlungsalternativen vorgegeben sind, was eine zusätzliche kognitive Anforderung an die Probanden darstellen sollte.

Als eine weitere Einschränkung dieser Studie ist zu verzeichnen, dass neben der Erfassung des lesedidaktischen Wissens (PCK) und PK kein Maß für CK eingesetzt wurde. Ursächlich hierfür ist, dass zum Zeitpunkt der Durchführung der Studie kein geeignetes, veröffentlichtes und standardisiertes Testinstrument zur Erfassung von CK bezogen auf den Leseunterricht zur Verfügung stand. Es konnte zwar hypothesenkonform ein positiver Zusammenhang zwischen dem lesedidaktischen Wissen und PK gezeigt werden, jedoch fehlt ein Maß zu CK, um die Struktur und Entwicklung des professionellen Wissens von Deutschlehramtsstudierenden umfassend untersuchen zu können. Im weiteren Projektverlauf des FuN-Kollegs EKoL wurde ein CK-Test für den Bereich der Lesedidaktik entwickelt für den eine umfassende Skalierung noch aussteht. So soll eine Untersuchung der Zusammenhänge zwischen PCK und CK bezogen auf den Leseunterricht sowie PK ermöglicht werden.

Die Ergebnisse zum Einfluss der praktischen Erfahrungen der Studierenden auf das lesedidaktische Wissen sollten ebenfalls mit gewissen Einschränkungen interpretiert werden: Sowohl die Variable „Lehrerfahrung im Leseunterricht“ als auch die Variable „Besuch lesedidaktischer Seminare“ wurden in der vorliegenden Studie dichotom erfasst. Weiterführend wäre wichtig zu überprüfen, ob auch die Quantität dieser Variablen das lesedidaktische Wissen (mit-)beeinflusst; es sollte

daher anschließend untersucht werden, ob das lesedidaktische Wissen mit der Anzahl an besuchten lesedidaktischen Seminaren ansteigt oder nicht. Des Weiteren ist aus anderen Studien bekannt, dass die Qualität von Lehrveranstaltungen zwischen und innerhalb unterschiedlicher Hochschulen stark variiert (Merzyn, 2002). Inwiefern die Qualität und Quantität von praktischen Erfahrungen auf die Entwicklung des lesedidaktischen Wissens wirkt, kann an dieser Stelle (noch) nicht abschließend und umfassend beantwortet werden.

Daneben ist zu beachten, dass es sich bei dem Ausbildungsbaustein „Lehrerfahrung“ bislang um eine „Black Box“ handelt. Es wird zwar angenommen, dass die eigene Lehrerfahrung der angehenden Lehrkräfte bedeutsam für den Aufbau professioneller Kompetenzen von Lehrkräften ist, allerdings liegen bislang keine Befunde darüber vor *was* es ist, das die Lehrerfahrung bedeutsam macht (Hascher, 2006). Es können verschiedene Entwicklungsprozesse angenommen werden, die durch die eigene Lehrerfahrung angestoßen und gefördert werden: Zunächst könnte das wiederholte Einüben von bestimmten Verhaltensweisen im Unterricht zu einer Steigerung der professionellen Kompetenz führen (Bollonow, 1991; Neuweg, 2005). Möglich ist auch, dass theoretisch erworbenes Wissen durch die Lehrerfahrung zum Tragen kommt und so ein Integrationsprozess von theoretischem Wissen und praktischem Können stattfinden kann (Hollins, 2011). Womöglich finden auch Entwicklungsprozesse bezüglich Überzeugungen (Behrmann & Souvignier, 2013) oder des Selbstkonzeptes (Retelsdorf et al., 2014b) statt, die zu einem verbesserten unterrichtlichen Handeln führen. Welche dieser Faktoren für den Auf- und Ausbau professioneller Kompetenzen von angehenden Lehrkräften insbesondere bedeutsam sind und inwiefern diese Entwicklungsprozesse interagieren, kann an dieser Stelle noch nicht beantwortet werden. Hierzu werden echte Längsschnittdaten benötigt, die Lehramtsstudierende bzw. Referendare über einen längeren Zeitraum hinweg untersuchen und zu mehreren Messzeitpunkten wiederkehrend verschiedene professionelle Kompetenzaspekte abfragen. Damit könnte beispielsweise empirisch überprüft werden, inwiefern das berufsbezogene Selbstkonzept der angehenden Lehrkräfte zu Beginn des Referendariats das Selbstkonzept am Ende des Referendariats sowie weitere kognitive und affektive professionelle Teilkompetenzen beeinflusst.

Zuletzt ist anzumerken, dass die Befragungen der Studierenden während bzw. anstatt regulärer Seminarsitzungen stattfanden. In der Regel waren die Befragungen unangekündigt. Als Belohnung erhielten die Studierenden zwar Süßigkeiten und konnten an einem Gewinnspiel teilnehmen, trotzdem ist – u. a. aufgrund der Bearbeitungsdauer von ca. 90 Minuten – fraglich, ob alle Studierenden den Test motiviert und aufmerksam bearbeitet haben. Es wäre interessant, ob die Ergebnisse ähnlich oder verschieden ausfallen, wenn die (extrinsische) Motivation der Studierenden durch einen entsprechenden (finanziellen) Anreiz oder durch den Einsatz des Testinstruments in einer Prüfungssituation erhöht werden würde. Eine höhere Motivation der Studierenden könnte über selbst eingeschätzte Motivationsskalen sowie eine Analyse über das Ausmaß an Missings im Fragebogen realisiert werden: Es würde hierbei erwartet, dass Studierende, die für das Bearbeiten des Vignetentests eine attraktive Belohnung erhalten, eine subjektiv höhere Motivation beim Bearbeiten des Fragebogens berichten und weniger Items auslassen als Studierende, die keine Belohnung erhalten.

14 Implikationen für die Praxis

Aus den Ergebnissen dieser Studie resultieren verschiedene Implikationen, die die Entwicklung von Testinstrumenten zur Erfassung von fachdidaktischen Wissensinhalten (angehender) Lehrkräfte, Handlungsempfehlungen für die Gestaltung der Lehramtsausbildung an der Hochschule sowie Anwendungsmöglichkeiten von Vignettentests betreffen:

Zunächst erscheint das Testformat des Vignettentests mit geschlossenem Antwortformat als geeignet, um professionelle Wissensinhalte mit impliziten Anteilen – wie im vorliegenden Fall das lesedidaktische Wissen von Lehramtsstudierenden – zu erfassen. Dieses spezifische Testformat hat sich daher auch für das geisteswissenschaftliche Unterrichtsfach Deutsch bewährt; bislang lagen ausschließlich gesicherte empirische Evidenzen über den Nutzen des Einsatzes von Vignettentests zur Erfassung von PCK in den Naturwissenschaften vor (Brovelli et al., 2014; Tepner & Dollny, 2014). Es erscheint deshalb lohnenswert auch für benachbarte geisteswissenschaftliche Domänen vignettenbasierte Testverfahren zur Erfassung und Exploration von PCK zu erstellen und einzusetzen. So kann der Nutzen und die Eignung dieser Testverfahren für die geisteswissenschaftlichen Unterrichtsfächer weiterführend untersucht und die empirische Befundlage ausgebaut werden.

Ein wichtiges Ergebnis der vorliegenden Studie ist, dass insbesondere die praktischen Erfahrungen der Deutschstudierenden für Zuwächse des lesedidaktischen Wissens im Laufe des Lehramtsstudiums verantwortlich sind. Unter praktischer Erfahrung wurde in dieser Studie einerseits die Lehrerfahrung verstanden, die Studierende bereits während ihres Studiums durch entsprechende semesterbegleitende Praktika erworben haben. Dieses Ergebnis schließt an Annahmen zur Entwicklung von PCK aus Studien im Kontext der naturwissenschaftlichen Unterrichtsfächer an: Hier wird postuliert, dass sich PCK substanziell durch die eigenen Erfahrungen der Lehramtsstudierenden im Unterrichten entwickelt (vgl. Lee et al., 2007). Ferner berichten Studien, dass der professionelle Wissensinhalt PCK nach einem semesterbegleitenden Schulpraktikum in einer stärker integrierten Form vorliegt, als vor dem Praktikum (Aydin, Demirdogen, Nur Akin, Uzuntiryaki-Kondakci & Tarkin, 2015). Auch berichten Lehramtsstudierende einen signifikan-

ten Anstieg hinsichtlich des selbsteingeschätzten Zuwachses an (Lehr-)Erfahrungen durch studienbegleitende Praktika (Choy, Wong, Goh & Ling Low, 2014). In längsschnittlichen Analysen zur Untersuchung der Kompetenzentwicklung im Referendariat könnte überprüft werden, inwiefern die selbsteingeschätzte fachliche, fachdidaktische und pädagogisch-psychologische Kompetenz mediierend bzw. moderierend auf die tatsächliche Kompetenzentwicklung – die über standardisierte Testverfahren erfasst wird – wirkt.

Die Rolle und Bedeutung von studienbegleitenden Praktika und deren adäquate Begleitung durch qualifizierte Mentoren sollten daher zukünftig noch stärker in den Fokus der Forschung zu Wirkfaktoren in der Lehramtsausbildung gestellt werden. Durch diese Studie konnte ferner der Status quo der Lehramtsausbildung an den Pädagogischen Hochschulen in Baden-Württemberg weiterführend (empirisch) unterstützt werden, deren Ausbildungskonzepte – u. a. durch obligatorische begleitete Semesterpraktika – es vorsehen, die theoretischen und praktischen Phasen der Lehramtsausbildung eng miteinander zu verknüpfen. Auf diese Art und Weise soll eine ganzheitliche Entwicklung von professionellen Wissensinhalten begünstigt werden (König, Tachtsoglou, Darge & Lünemann, 2014).

Andererseits legen die Ergebnisse dieser Studie nahe, dass neben der eigenen Lehrerfahrung die Nutzung spezifischer Lerngelegenheiten, wie der Besuch lesedidaktischer Seminare, als bedeutsam für die Entwicklung des lesedidaktischen Wissens einzuschätzen ist. Ähnliche Ergebnisse wurden in einer Studie zur Schreibdidaktik erzielt: Durch den Besuch eines spezifischen Seminarangebotes zu Schreibtechniken entwickelten die teilnehmenden Studierenden (im Selbstbericht) zunehmend das Rollenverständnis einer Lehrkraft, die kompetent Schreibprozesse anleiten kann (Martin & Dismuke, 2015). In Bezug auf angehende Mathematiklehrkräfte im internationalen Vergleich liegen Evidenzen vor, dass spezifische Lerngelegenheiten im Studium auf das mathematikdidaktische Wissen wirken (Blömeke et al., 2010). Hier konnte allerdings nur ein positiver Einfluss von den Lerngelegenheiten in Mathematik auf das mathematikdidaktische Wissen gezeigt werden, nicht jedoch von den Lerngelegenheiten in der Mathematikdidaktik auf das mathematikdidaktische Wissen. Die Autoren nehmen hierfür u. a. methodische Gründe an. Auch von Seiten der Studierenden werden immer wieder praxisrelevante und schulnahe Seminare und Veranstaltungen eingefordert, die be-

reits zu Beginn des Studiums angeboten werden sollten (Zaunbauer, Brouer, Schmidt & Möller, 2015). Unklar ist, über welches Ausmaß an CK die Lehramtsstudierenden verfügen sollten bzw. müssen, um nachhaltig bezüglich PCK gefördert werden zu können. Hier bietet sich die Annahme eines Schwellenmodells an: Lehramtsstudierende müssen über ein gewisses Ausmaß an CK verfügen; wenn diese Schwelle überschritten ist, kann sich darauf aufbauend PCK entwickeln (Baumert & Kunter, 2006). Eine offene Frage ist, wie hoch diese „CK-Schwelle“ ist und ob sich diese Schwellen für unterschiedliche Fachdidaktiken unterscheiden.

Kurzum sollten aufgrund der empirischen Evidenzen zum Nutzen fachdidaktischer Veranstaltungen für die Entwicklung von PCK und dem allgemeinen Interesse der Studierenden an solchen Veranstaltungen, vielfältige (fach-)didaktische Veranstaltungen über die gesamte Studiendauer hinweg angeboten werden. Dies trägt dann voraussichtlich nicht nur zur Steigerung der professionellen Wissensinhalte der Lehramtsstudierenden bei, sondern auch zu deren Zufriedenheit mit den Studieninhalten.

Dass die Lehrerfahrung im Leseunterricht sowie der Besuch lesedidaktischer Seminare prädiktiv für das lesedidaktische Wissen der Studierenden sind, stützt die Grundannahme der vorliegenden Studie, dass durch die Verwendung eines Vignetentests mit geschlossenem Antwortformat (auch) implizite Anteile des lesedidaktischen Wissens erfasst werden. Unter der Annahme, dass PCK ein professioneller Wissensinhalt mit impliziten und expliziten Anteilen ist, ergeben sich Implikationen für den Wissenstransfer an der Hochschule; es kann nämlich davon ausgegangen werden, dass sich die Vermittlung impliziter Wissensinhalte – insofern dies möglich ist – von der Vermittlung expliziter Wissensinhalte unterscheidet.

In diese Überlegung können sinnvollerweise Modelle aus der Unternehmensforschung miteinbezogen werden. Auch hier beschäftigt man sich mit Möglichkeiten der Vermittlung von impliziten Wissensinhalten.

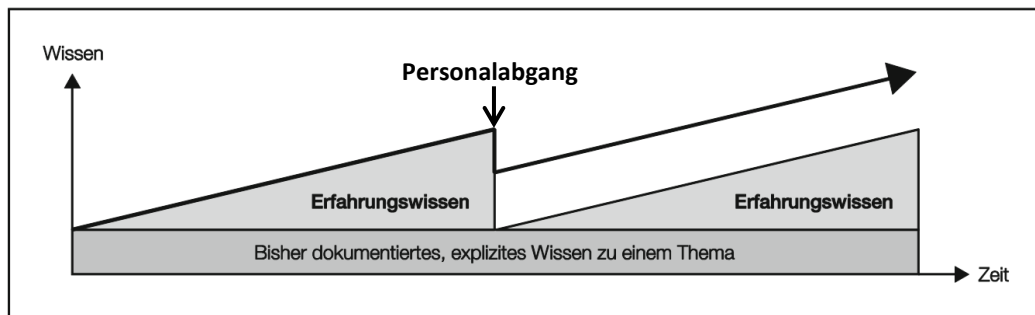


Abbildung 25. Wissensverlust bei Personalabgängen (aus: Sutter, 2016, S. 22).

In Abbildung 25 wird dargestellt, dass es bei Personalabgängen von langjährigen Mitarbeitern zu einem Verlust an unternehmens- und berufsbezogenem Erfahrungswissen kommt. Der ausscheidende und der neu eingestellte Beschäftigte verfügen über eine gemeinsame Wissensbasis an explizitem Wissen, das u. a. in der Ausbildung erworben wurde. Allerdings hat der ausscheidende Beschäftigte durch eigene Erfahrungen zusätzlich implizites Wissen aufgebaut („Erfahrungswissen“ laut Abbildung), das mit zunehmender Beschäftigungsdauer im Betrieb linear angestiegen ist. Für das Wissensmanagement im Unternehmen ist außerordentlich wichtig, dass dieses implizite Erfahrungswissen von langjährigen Mitarbeitern nicht durch deren Ausscheiden aus dem Unternehmen verloren geht, sondern dass dieses – zumindest teilweise – an den nachfolgenden Beschäftigten weitergegeben wird. Um dieses Erfahrungswissen möglichst umfassend weiterzugeben, werden bezogen auf Unternehmen u. a. Methoden zur graphischen Visualisierung wie Mind-Maps oder Concept-Maps als Transfermodell für implizites Wissen diskutiert (vgl. Sutter, 2016).

Übertragen auf den Kontext der Lehramtsausbildung ist es bezüglich der Planung und der Gestaltung des Lehrangebots wichtig zu bedenken, dass das professionelle Wissen von Lehrkräften implizite sowie explizite Anteile umfasst. Um implizites Wissen an Dritte weitervermitteln zu können ist es notwendig, dieses zumindest zu einem gewissen Grad zu explizieren (Sutter, 2016). Wie könnte die Vermittlung von impliziten Anteilen des professionellen Wissens im Rahmen von Lehrveranstaltungen im Lehramtsstudium ermöglicht werden?

Es wäre denkbar, fachdidaktische Problemstellungen am Modell zu besprechen und zu üben, beispielsweise im Rollenspiel oder mithilfe von E-Learning-

Programmen (Gartmeier et al., 2015; Mirliss, 2014; Rao & Stupans, 2012). Die Studien zu Effekten von Rollenspielen und E-Learning-Programmen zielen bislang allerdings hauptsächlich auf eine Steigerung der kommunikativen Kompetenzen der angehenden Lehrkräfte ab. Hier werden positive Effekte berichtet (Gartmeier et al., 2015). Außerdem könnten umfassende Mentoring- bzw. Tutoring-Angebote, die eine Möglichkeit zur Nachbesprechung bieten (z. B. der eigenen Lehrerfahrung im semesterbegleitenden Praktikum), zur Explizierung impliziter (Lern-)Prozesse hilfreich sein (für einen Überblick zu Mentoring: Johnson & Ridley, 2008).

Daneben wird aktuell der Nutzen des Einsatzes von Unterrichtsvideos in der Lehramtsausbildung diskutiert (Blomberg, Renkl, Sherin, Borko & Seidel, 2013; Borko, Koellner, Jacobs & Seago, 2011; Kleinknecht & Schneider, 2013; Sherin & van Es, 2008). Unterrichtsvideos sind kurze Aufnahmen von Lehrer-Schüler-Interaktionen aus dem Schulunterricht; diese können realistisch oder nachgestellt sein. Es wird angenommen, dass durch das Beobachten von Unterrichtsvideos wichtige professionelle Kompetenzen – z. B. die professionelle Wahrnehmung von Unterrichtssituationen – aufgebaut werden, die es den Studierenden ermöglichen, adäquate Handlungsmöglichkeiten für die beobachteten Unterrichtssequenzen abzuleiten (Borko et al., 2011). Unterrichtsvideos sollen es den Studierenden ermöglichen, sich in die Lehrsituationen hineinzusetzen und durch die Perspektivenübernahme eine „secondhand experience“ zu erleben (vgl. Kleinknecht & Schneider, 2013). Das bedeutet, dass sich den Lehramtsstudierenden durch die Analyse und Reflexion von Unterrichtsvideos ein „Fenster zur Lehrpraxis“ öffnet, ohne dass diese selbst mit Schülern interagieren (müssen) (Blomberg et al., 2013). In dieser Studie wurde auf die Verwendung von Unterrichtsvideos verzichtet, da sich für die Erfassung von kognitiven Konstrukten – wie dem lesedidaktischen Wissen – textbasierte Vignetten besonders eignen (Herbst & Kosko, 2014). Videovignetten eignen sich in besonderem Maße zur Erfassung von Konstrukten der professionellen Wahrnehmung wie z. B. der Professionellen Unterrichtswahrnehmung (englisch: *professional vision*) (Meschede, Steffensky, Wolters & Möller, 2015).

Der Einsatz von Rollenspielen, E-Learning-Programmen und Unterrichtsvideos erscheint daher als vielversprechende Möglichkeit – sozusagen durch eine „Vor-

wegnahme – von eigener Lehrerfahrung, implizite Wissensinhalte von Lehramtsstudierenden aufzubauen. Als besonders wertvoll für den Aufbau impliziter Wissensinhalte wird die ausführliche, durch den Dozenten moderierte Reflexion der betrachteten Unterrichtssituationen bzw. des Rollenspiels erachtet. In einer längsschnittlichen Untersuchung von drei angehenden Chemielehrkräften wurden Mentoring bezüglich PCK, die eigene Lehrerfahrung sowie die Beobachtung von anderen angehenden Lehrkräften beim Unterrichten als bedeutsam für den Aufbau professioneller Kompetenz herausgestellt (Aydin et al., 2013). Angelehnt daran liegen erste empirische Evidenzen vor, dass implizite Wissensinhalte bereits während des Studiums aufgebaut werden können, wenn theoretisches Wissen kontextualisiert und auf spezifische Problemstellungen bezogen vermittelt wird (Collin & Tynjälä, 2003).

Zuletzt soll eine weitere Einsatzmöglichkeit des Testformates Vignettentest abgeleitet werden: Vignettentests könnten zukünftig zur Eignungsfeststellung für das Lehramtsstudium bzw. als Selbsterkundungsverfahren herangezogen werden, um die angehenden Lehramtsstudierenden bezüglich ihrer Studienwahl zu unterstützen.

Bislang gibt es bundesweit keine Hochschule, die ein Eignungsfeststellungsverfahren für Lehramtsstudiengänge anwendet (Bremerich-Vos & Dämmer, 2013). Allerdings sind Online-Self-Assessments bezüglich (persönlicher) Einstellungen der zukünftigen Lehramtsstudierenden bereits an vielen Hochschulen obligatorisch. Da es kaum geeignete Instrumente gibt, um fachliche Voraussetzungen von angehenden Lehramtsstudierenden zu überprüfen, sind diese Online-Self-Assessments in der Regel auf Persönlichkeitseigenschaften und Einstellungen beschränkt (Rothland & Tirre, 2011). Oser (2006) fordert, dass bei der Auswahl geeigneter Lehramtsstudierender neben motivationalen Aspekten und Einstellungen auch professionelle Wissensinhalte wie die „fachdidaktische Kompetenz“ berücksichtigt werden sollten.

Ein Einsatz von vignettenbasierten Testverfahren würde ein Eignungsfeststellungsverfahren von angehenden Lehramtsstudierenden daher ergänzen: Durch die Bearbeitung der Vignetten kann der momentane Wissensstand über fachdidaktische Inhalte des Faches abgebildet werden. Es ist anzunehmen, dass sich die angehenden Lehramtsstudierenden bezüglich des fachdidaktischen Wissens und

Könnens bereits vor Beginn des Studiums unterscheiden: Es liegen hierzu Befunde vor, dass das Ausmaß an Lehrerfahrung, die Lehramtsstudierende vor Aufnahme ihres Studiums erworben haben, einen positiven Zusammenhang mit der späteren Leistung in praxisbezogenen Aufgaben („teaching practice grades“) aufweist (Heinz, 2013). Daneben kann ein Vignettentest als zusätzlicher Baustein eines Selbsterkundungsverfahrens/Self-Assessments für Lehramtsstudierende verstanden werden: Durch die Bearbeitung der Unterrichtsvignetten könnten die Studierenden die Perspektive einer Lehrkraft einnehmen und einen realitätsnahen Einblick in die Anforderungen und Aufgaben erhalten, die der Lehrerberuf mit sich bringt.

Der Einsatz von vignettenbasierten Testinstrumenten könnte daher dazu beitragen, möglichst geeignete Kandidaten für die Lehramtsausbildung auszuwählen. Des Weiteren könnte auf diese Art und Weise angehenden Lehramtsstudierenden bei ihrer Berufswahl ein Anhaltspunkt zur kritischen Reflexion der eigenen Eignung und Motivation geboten werden.

15 Ausblick über anschließende Forschungsfragen

Nachdem die Ergebnisse dieser Dissertationsstudie sowie Einschränkungen bezüglich der Interpretation diskutiert und entsprechende Implikationen abgeleitet und erläutert wurden, soll in diesem Abschnitt ein Ausblick auf weitere Forschungsfragen gegeben werden, die anschließend an diese Studie bearbeitet werden sollen.

Zunächst soll die Retest-Reliabilität des Vignettentests überprüft werden, um das Ergebnis zur EAP-Reliabilität des Testverfahrens zu ergänzen und zu präzisieren. Die Bestimmung der Retest-Reliabilität ist besonders angemessen für Testinstrumente, die heterogene Konstrukte erfassen (Lienert & Raatz, 1998). Aufgrund der Testanforderung (Einschätzung von vier bis sechs geschlossenen Items pro Vignette auf einer sechsstufigen Likert-Skala) ist nicht von Erinnerungs- oder Lerneffekten durch die Testbearbeitung auszugehen. Zudem ist zu erwarten, dass das lesedidaktische Wissen von Lehramtsstudierenden über kurze Zeiträume stabil bleibt (außer die Studierenden nehmen an spezifischen Lerngelegenheiten teil oder erwerben Lehrerfahrung). Daher erscheint ein Retest-Intervall von 4 Wochen als angemessen, um die Retest-Reliabilität des Vignettentests zu untersuchen. Eine allgemein gültige Empfehlung über die Wahl des Retest-Intervalls kann allerdings nicht gegeben werden, da diese immer vom jeweiligen Testinhalt abhängt (Moosbrugger & Kelava, 2012).

In weiteren Erhebungen soll versucht werden, schwierige sowie leichte Vignetten in das Testinstrument zu integrieren, sodass ein ausgewogenes Verhältnis bezüglich der Schwierigkeit der Testaufgaben resultiert. Hierzu sollen Vignetten verwendet werden, die im Rahmen der Testentwicklung erstellt, dann jedoch nicht für das Testinstrument verwendet wurden (19 Vignetten). Geplant ist, zwei schwierige und zwei leichte Vignetten in das Testinstrument zu integrieren. Dazu sollen nochmals qualitative Expertenbefragungen durchgeführt werden. Durch dieses Vorgehen soll eine gesteigerte Differenzierungsfähigkeit des Testinstru-

ments im oberen und unteren Bereich der Skala erreicht werden (Moosbrugger & Kelava, 2012).

Ein weiteres Anliegen anschließender Forschung ist, den Vignettentest zur Erfassung impliziter Anteile des lesedidaktischen Wissens gemeinsam mit etablierten Testverfahren einzusetzen, die vorrangig die expliziten Anteile des lesedidaktischen Wissens fokussieren (vgl. Bremerich-Vos & Dämmer, 2013; Pissarek & Schilcher, 2015). So könnte z. B. untersucht werden, ob sich die beiden Wissensdimensionen explizites und implizites lesedidaktisches Wissen empirisch abbilden lassen. Zudem könnten so Hinweise auf die konvergente Validität des Testverfahrens gewonnen werden (Moosbrugger & Kelava, 2012).

Weiterhin soll die Querschnitts-Stichprobe von Deutschstudierenden und Referendaren mit dem Fach Deutsch um erfahrene Lehrkräfte erweitert werden. So soll überprüft werden, ob das lesedidaktische Wissen nach dem Studium noch weiter „anstiegt“. Dies könnte als weiterer Beleg für die Validität des Vignettentests gewertet werden. Daran anschließend soll ein Kontrollgruppenvergleich auf „Experten-ebene“ durchgeführt werden (vgl. Jütter & Neuhaus, 2013): Dazu soll der Vignettentest neben erfahrenen Lehrkräften auch Germanisten/Sprachwissenschaftlern sowie erfahrenen Pädagogen vorgelegt werden. Hierbei werden deutlichere Unterschiede zwischen den Gruppen erwartet als bei dem Kontrollgruppenvergleich auf der Ebene von Studierenden, der in dieser Studie berichtet wurde.

Ferner soll im Anschluss an diese Studie der Einfluss der Lehrerfahrung der Studierenden auf das lesedidaktische Wissen dezidiert untersucht werden. Hierzu soll eine Prä-Post- und Follow-Up Messung von Deutschstudierenden im ISP (Integriertes Semesterpraktikum) bzw. Referendariat mit entsprechenden Kontrollgruppen durchgeführt werden. So soll überprüft werden, inwiefern das ISP bzw. das Referendariat zu einer individuellen Steigerung des lesedidaktischen Wissens der Lehramtsstudierenden bzw. der Referendare führt. Auf diese Art und Weise kann eine echte längsschnittliche Untersuchung der Entwicklung der professionellen Kompetenz im Verlauf der praktischen Ausbildungsabschnitte im Lehramtsstudium realisiert werden.

Zusätzlich wird angedacht in nachfolgenden Studien einen Indikator für die Unterrichtsqualität einzusetzen. König, Buchholtz und Dohmen (2015) haben die Planungskompetenz von Lehramtsstudierenden als einen wichtigen Indikator für die Unterrichtsqualität untersucht. Die Entwicklung der Planungskompetenz stellt ein wichtiges Ziel der Lehramtsausbildung dar und gilt als allgemeine didaktische Fähigkeit, die essenziell für die erfolgreiche Gestaltung des Unterrichts ist. König et al. (2015) haben ein Verfahren vorgestellt, das Unterrichtsentwürfe von Referendaren anhand von festgelegten Kriterien kodiert. Die IRT-Skalierung des Verfahrens zeigt zufriedenstellende Ergebnisse, sodass dieses Verfahren als ein reliables Maß für die Planungskompetenz angehender Lehrkräfte angesehen werden kann. Durch den gemeinsamen Einsatz des Vignettentests mit dem Verfahren zur Überprüfung der Planungskompetenz als Indikator für Unterrichtsqualität könnte näherungsweise untersucht werden, inwiefern sich der professionelle Wissensinhalt lesedidaktisches Wissen im professionellen Handeln im Unterricht niederschlägt.

Das Fernziel – nach umfangreichen Untersuchungen zur Reliabilität und weiteren Aspekten der Validität des Testverfahrens – stellt der Einsatz des Vignettentests im Klassenzimmer dar: Letztendlich soll überprüft werden, ob das lesedidaktische Wissen von Lehrkräften die Leistungsentwicklung der Schüler kausal bedingt. Um eine solche Fragestellung zu untersuchen, bietet sich ein Cross-Lagged-Panel Design an, das zur Analyse von kausalen Fragestellungen geeignet ist (Reinders & Heinz, 2006). Zusammengefasst müssten in einem Cross-Lagged-Panel Design zu Beginn eines Schuljahres das lesedidaktische Wissen der Lehrkräfte und die Leseleistung der Schüler, z. B. mit dem Testinstrument LGVT 6-12 (Lesegeschwindigkeits- und -verständnistest für die Klassen 6-12) erfasst werden (Schneider, Schlagmüller & Ennemoser, 2007). Daneben würden eine Reihe an Kovariaten erhoben werden, um motivationale Aspekte, Einstellungen, Persönlichkeitseigenschaften etc. der Lehrer und Schüler kontrollieren zu können. Am Ende des Schuljahres würde die Erhebung ein zweites Mal durchgeführt: Wieder wird das lesedidaktische Wissen der Lehrkräfte (das ohne spezifische Weiterbildung konstant bleiben sollte) sowie die Leseleistung der Schüler (die in Abhängigkeit des professionellen Wissens der Lehrkraft angestiegen sein sollte) getestet. Durch die Verwendung eines Cross-Lagged-Panel Designs kann überprüft werden, inwiefern

ein höheres Ausmaß an lesedidaktischem Wissen auf Seiten der Lehrkraft die Leistungsentwicklung der Schüler begünstigt. Auf diese Art und Weise könnte der (kausale) Einfluss des lesedidaktischen Wissens von Lehrkräften auf die Schülerleistung evidenzbasiert untersucht werden.

16 Fazit

Im Rahmen dieser Dissertationsstudie wurde ein vignettenbasiertes Testinstrument mit geschlossenem Antwortformat zur Erfassung der impliziten Anteile des fachdidaktischen Wissens im Leseunterricht bei angehenden Lehrkräften entwickelt. Die erzielten Ergebnisse zur Validität und Reliabilität können als vielversprechend bezeichnet werden. Das Testformat „Vignettentest mit geschlossenem Antwortformat“ hat sich daher als geeignet für das geisteswissenschaftliche Unterrichtsfach Deutsch erwiesen.

Die Befunde dieser Studie weisen darauf hin, dass durch den Vignettentest tatsächlich implizite Anteile des lesedidaktischen Wissens erfasst werden (z. B. durch die Abhängigkeit des lesedidaktischen Wissens von der spezifischen Lehrerfahrung im Leseunterricht). Inwiefern sich das lesedidaktische Wissen von (angehenden) Lehrkräften im professionellen Handeln niederschlägt und ob die Leistungsentwicklung der Schüler hierdurch (mit-)beeinflusst wird, müssen anschließende Forschungsarbeiten klären.

Zuletzt ergeben sich durch diese Studie vielfältige Implikationen, die sowohl die Entwicklung und den Einsatz standardisierter vignettenbasierter Testinstrumente als auch Handlungsempfehlungen für die Qualitätssicherung in der Lehramtsausbildung an der Hochschule betreffen. Diese Studie kann daher als fundierter und wertvoller Ansatzpunkt zum besseren Verständnis und der Optimierung von Lehr-Lernprozessen bezüglich des fachdidaktischen Wissens im Leseunterricht im Rahmen der universitären Hochschulbildung gewertet werden.

17 Literaturverzeichnis

- Abd-El-Khalick, F. (2006). Preservice and experienced biology teachers' global and specific subject matter structures: Implications for conceptions of pedagogical content knowledge. *Eurasia Journal of Mathematics, Science and Technology Education* (2), 1-29.
- Abele, A. E. (2011). Prädiktoren des Berufserfolgs von Lehrkräften. Befunde der Langzeitstudie MATHE. *Zeitschrift für Pädagogik*, 57, 674-694.
- Abele, A. E. & Candova, A. (2007). Prädiktoren des Belastungserlebens im Lehrerberuf. *Zeitschrift für Pädagogische Psychologie*, 21 (2), 107-118.
- Adams, R. & Wu, M. L. (2002). PISA 2000 Technical Report.
- Adams, R. J., Wu, M. L. & Wilson, M. (2012). The Rasch Rating Model and the Disordered Threshold Controversy. *Educational and Psychological Measurement*, 72 (4), 547-573.
- Albisser, S., Kirchhoff, E. & Albisser, E. (2009). Berufsmotivation und Selbstregulation: Kompetenzentwicklung und Belastungserleben von Studierenden, berufseinsteigenden und erfahrenen Lehrpersonen. *Unterrichtswissenschaft*, 37 (3), 262-288.
- Alisch, L.-M., Hermkes, R. & Möbius, K. (2009). Messen von Lehrprofessionalität 1: Grundlagen. In O. Zlatkin-Troitschanskaia (Hrsg.), *Lehrprofessionalität. Bedingungen, Genese, Wirkungen und ihre Messung* (Beltz-Bibliothek, S. 249-262). Weinheim: Beltz.
- Allison, P. D. (2002). *Missing Data*. Thousand Oaks: Sage Publications.
- Alonzo, A. C. & Kim, J. (2015). Declarative and dynamic pedagogical content knowledge as elicited through two video-based interview methods. *Journal of Research in Science Teaching*, 1-28.
- Ames, A. J. & Penfield, R. D. (2015). An NCME Instructional Module on Item-Fit Statistics for Item Response Theory Models. *Educational Measurement: Issues and Practice*, 34 (3), 39-48.

- Arens, A. K. & Morin, A. J. S. (2016). Relations Between Teachers' Emotional Exhaustion and Students' Educational Outcomes. *Journal of Educational Psychology*.
- Artelt, C. (2009). Diagnostische Urteile von Lehrkräften im Bereich der Lesekompetenz. In A. Bertschi-Kaufmann & C. Rosebrock (Hrsg.), *Literalität. Bildungsaufgabe und Forschungsfeld* (Lesesozialisation und Medien, 125-136). Weinheim: Juventa-Verl.
- Artelt, C., Beinicke, A., Schlagmüller, M. & Schneider, W. (2009). Diagnose von Strategiewissen beim Textverstehen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 41 (2), 96-103.
- Artelt, C. & Dörfler, T. (2010). Förderung von Lesekompetenz als Aufgabe aller Fächer. Forschungsergebnisse und Anregungen für die Praxis. In H. Ruch (Hrsg.), *ProLesen - auf dem Weg zur Leseschule* (Leseförderung in den gesellschaftswissenschaftlichen Fächern, Aufsätze und Materialien aus dem KMK-Projekt "ProLesen", S. 13-36). Donauwörth: Auer.
- Artelt, C. & Gräsel, C. (2009). Diagnostische Kompetenz von Lehrkräften. *Zeitschrift für Pädagogische Psychologie*, 23 (34), 157-160.
- Artelt, C., Naumann, J. & Schneider, W. (2010). Lesemotivation und Lernstrategien. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel et al. (Hrsg.), *PISA 2009. Bilanz nach einem Jahrzehnt* (S. 73-112). Münster: Waxmann.
- Artelt, C., Stanat, P., Schneider, W. & Schiefele, U. (2001). Lesekompetenz: Testkonzeption und Ergebnisse. In J. Baumert, E. Klieme, M. Neubrand, M. Prenzel, U. Schiefele, W. Schneider et al. (Hrsg.), *PISA 2000* (Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich, S. 69-137). Opladen: Leske + Budrich.
- Ayala, R. J. de. (2009). *The theory and practice of item response theory* (Methodology in the social sciences). New York: Guilford Press.
- Aydeniz, M. & Kirbulut, Z. D. (2014). Exploring challenges of assessing pre-service science teachers' pedagogical content knowledge (PCK). *Asia-Pacific Journal of Teacher Education*, 42 (2), 147-166.

- Aydin, S., Demirdogen, B., Nur Akin, F., Uzuntiryaki-Kondakci, E. & Tarkin, A. (2015). The nature and development of interaction among components of pedagogical content knowledge in practicum. *Teaching and Teacher Education*, 46, 37-50.
- Aydin, S., Demirdogen, B., Tarkin, A., Kutucu, S., Ekiz, B., Nur Akin, F. et al. (2013). Providing a Set of Research-Based Practices to Support Preservice Teachers' Long-Term Professional Development as Learners of Science Teaching. *Science Education*, 97 (6), 903-935.
- Backhaus, K., Erichson, B. & Weiber, R. (Hrsg.). (2015). *Fortgeschrittene Multivariate Analysemethoden. Eine anwendungsorientierte Einführung* (3., überarbeitete und aktualisierte Auflage). Berlin, Heidelberg: Springer Gabler.
- Baer, M., Dörr, G., Fraefel, U., Kocher, M., Küster, O., Larcher, S. et al. (2007). Werden angehende Lehrpersonen durch das Studium kompetenter? Kompetenzaufbau und Standarderreichung in der berufswissenschaftlichen Ausbildung an drei Pädagogischen Hochschulen in der Schweiz und in Deutschland. *Unterrichtswissenschaft*, 35 (1), 15-47.
- Ball, D. L., Thames, M. H. & Phelps, G. (2008). Content Knowledge for Teaching: What Makes It Special? *Journal of Teacher Education*, 59, 389-407.
- Baron, R. M. & Kenny, D. A. (1986). The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations. *Journal of Personality and Social Psychology*, 51 (6), 1173-1182.
- Baumert, J., Blum, W. & Neubrand, M. (2004). Drawing the lessons from PISA 2000. *Zeitschrift für Erziehungswissenschaft*, 7 (3), 143-157.
- Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften [Keyword: Teachers' professional competence]. *Zeitschrift für Erziehungswissenschaft*, 4, 469-520.
- Baumert, J. & Kunter, M. (2011a). Das Kompetenzmodell von COACTIV. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 29-53). Münster [u.a.]: Waxmann.

- Baumert, J. & Kunter, M. (2011b). Das mathematikspezifische Wissen von Lehrkräften, kognitive Aktivierung im Unterricht und Lernfortschritte von Schülerinnen und Schülern. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 163-192). Münster [u.a.]: Waxmann.
- Baumert, J. & Kunter, M. (2013). Professionelle Kompetenz von Lehrkräften. In I. Gogolin, H. Kuper, H.-H. Krüger & J. Baumert (Hrsg.), *Stichwort: Zeitschrift für Erziehungswissenschaft* (S. 277-337). Wiesbaden: Springer Fachmedien Wiesbaden.
- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A. et al. (2010). Teachers' Mathematical Knowledge, Cognitive Activation in the Classroom, and Student Progress. *American Educational Research Journal*, 47 (1), 133-180.
- Behrmann, L. & Souvignier, E. (2013). Pedagogical content beliefs about reading instruction and their relation to gains in student achievement. *European Journal of Psychology of Education*, 28 (3), 1023-1044.
- Behrmann, L. & Souvignier, E. (2015). Effects of fit between teachers' instructional beliefs and didactical principles of reading programs. *European Journal of Psychology of Education*, 30 (3), 295-312.
- Berliner, D. C. (1988). *The Development of Expertise in Pedagogy*. Washington, D.C: AACTE Publications.
- Berry, C. J., Shanks, D. R. & Henson, R. N. A. (2008). A unitary signal-detection model of implicit and explicit memory. *Trends in cognitive sciences*, 12 (10), 367-373.
- Besser, M. & Krauss, S. (2009). Zur Professionalität als Expertise. In R. Mulder, O. Zlatkin-Troitschanskaia, K. Beck, R. Nickolaus & D. Sembill (Hrsg.), *Professionalität von Lehrenden - Zum Stand der Forschung* (S. 71-82). Weinheim: Beltz.

- Blomberg, G., Renkl, A., Sherin, M. G., Borko, H. & Seidel, T. (2013). Five research-based heuristics for using video in pre-service teacher education. *Journal for Educational Research Online*, 5 (1), 90-114.
- Blömeke, S., Bremerich-Vos, A., Haudeck, H., Kaiser, G., Nold, G., Schwippert, K. et al. (Hrsg.). (2011). *Kompetenzen von Lehramtsstudierenden in gering strukturierten Domänen. Erste Ergebnisse aus TEDS-LT*. Münster: Waxmann.
- Blömeke, S., Bremerich-Vos, A., Kaiser, G., Nold, G., Haudeck, H., Keßler, J.-U. et al. (Hrsg.). (2013). *Professionelle Kompetenzen im Studienverlauf Weitere Ergebnisse zur Deutsch-, Englisch- und Mathematiklehrerausbildung aus TEDS-LT*. Münster [u.a.]: Waxmann.
- Blömeke, S. & Delaney, S. (2012). Assessment of teacher knowledge across countries. A review of the state of research. *ZDM*, 44 (3), 223-247.
- Blömeke, S., Felbrich, A. & Müller, C. (2008). Erziehungswissenschaftliches Wissen am Ende der Lehrerausbildung. In S. Blömeke, G. Kaiser & R. Lehmann (Hrsg.), *Professionelle Kompetenz angehender Lehrerinnen und Lehrer: Wissen, Überzeugungen und Lerngelegenheiten deutscher Mathematikstudierender und -referendare - Erste Ergebnisse zur Wirksamkeit der Lehrerausbildung* (S. 195-217). Münster: Waxmann.
- Blömeke, S., Kaiser, G. & Lehmann, R. (Hrsg.). (2008). *Professionelle Kompetenz angehender Lehrerinnen und Lehrer: Wissen, Überzeugungen und Lerngelegenheiten deutscher Mathematikstudierender und -referendare - Erste Ergebnisse zur Wirksamkeit der Lehrerausbildung*. Münster: Waxmann.
- Blömeke, S., König, J., Suhl, U., Hoth, J. & Döhrmann, M. (2015). Wie situationsbezogen ist die Kompetenz von Lehrkräften? *Zeitschrift für Pädagogik*, 61 (3), 310-327.
- Blömeke, S., Suhl, U., Kaiser, G., Felbrich, A., Schmotz, C. & Lehmann, R. (2010). Lerngelegenheiten und Kompetenzerwerb angehender Mathematiklehrkräfte im internationalen Vergleich. *Unterrichtswissenschaft*, 38 (1), 29-50.
- Bollonow, O. F. (1991). *Vom Geist des Übens : Eine Rückbesinnung auf elementare didaktische Erfahrung* (3. Aufl.). Rothenhäusler: Stäfa.

- Bonnet, A. & Henricks, U. (2014). Professionalisierung und Deprofessionalisierung im Lehrer/innenberuf. Ansätze und Befunde aktueller empirischer Forschung. *Zeitschrift für interpretative Schul- und Unterrichtsforschung*, 3 (1), 3-12.
- Bonsen, M. & Rolff, H.-G. (2006). Professionelle Lerngemeinschaften von Lehrerinnen und Lehrern. *Zeitschrift für Pädagogik*, 52, 167-184.
- Borko, H., Koellner, K., Jacobs, J. & Seago, N. (2011). Using video representations of teaching in practice-based professional development programs. *ZDM*, 43 (1), 175-187.
- Borowski, A. & Riese, J. (2010). Physikalisch-fachdidaktisches Wissen: Was kommt in der Praxis an? *Praxis der Naturwissenschaften - Physik in der Schule*, 5 (95), 5-8.
- Bos, W., Tarelli, I., Bremerich-Vos, A. & Schwippert, K. (Hrsg.). (2012). *IGLU 2011: Lesekompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich*. Münster/New York/München/Berlin: Waxmann.
- Bos, W., Wendt, H., Köller, O. & Selter, C. (Hrsg.). (2012). *TIMSS 2011: Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern im internationalen Vergleich*. Münster/New York/München/Berlin: Waxmann.
- Bräuer, C. (2010). *Könnerschaft und Kompetenz in der Leseausbildung. Theoretische und empirische Perspektiven* (Lesesozialisation und Medien). Weinheim: Juventa-Verl.
- Bremerich-Vos, A. & Dämmer, J. (2013). Professionelles Wissen im Studienverlauf: Lehramt Deutsch. In S. Blömeke, A. Bremerich-Vos, G. Kaiser, G. Nold, H. Haudeck, J.-U. Keßler et al. (Hrsg.), *Professionelle Kompetenzen im Studienverlauf Weitere Ergebnisse zur Deutsch-, Englisch- und Mathematiklehrausbildung aus TEDS-LT* (S. 47-72). Münster [u.a.]: Waxmann.
- Bromme, R. (1992). *Der Lehrer als Experte*. Bern: Hans Huber.
- Bromme, R. (2014). *Der Lehrer als Experte. Zur Psychologie des professionellen Wissens* (Standardwerke aus Psychologie und Pädagogik, Reprints, Bd. 7, [Nachdr. der Ausg.] Bern 1992). Münster: Waxmann.

- Brophy, J. & Good, T. L. (1986). Teacher Behavior and Student Achievement. In M. C. Wittrock (Ed.), *Handbook of research on teaching. A project of the American Educational Research Association* (3rd ed., pp. 328-375). New York: Macmillan.
- Brovelli, D., Bölsterli, K., Rehm, M. & Wilhelm, M. (2013). Erfassen professioneller Kompetenzen für den naturwissenschaftlichen Unterricht - ein Vignettest mit authentisch komplexen Unterrichtssituationen und offenem Antwortformat. *Unterrichtswissenschaft*, 41, 306-329.
- Brovelli, D., Bölsterli, K., Rehm, M. & Wilhelm, M. (2014). Using Vignette Testing to Measure Student Science Teachers' Professional Competencies. *American Journal of Educational Research*, 2 (7), 555-558.
- Buchholz, N., Scheiner, T., Döhrmann, M., Suhl, U., Kaiser, G. & Blömeke, S. (2012). *TEDS-shortM. Kurzfassung der mathematischen und mathematikdidaktischen Testinstrumente aus TEDS-M, TEDS-LT und TEDS-Telekom*: Universität Hamburg.
- Bühner, M. (2004). *Einführung in die Test- und Fragebogenkonstruktion* (Psychologie). München [u.a.]: Pearson Studium.
- Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion* (Psychologie, 3., aktualisierte und erw. Aufl.). München: Pearson Studium.
- Carretti, B., Caldarola, N., Tencati, C. & Cornoldi, C. (2014). Improving reading comprehension in reading and listening settings: the effect of two training programmes focusing on metacognition and working memory. *The British journal of educational psychology*, 84, 194-210.
- Chan, K.-W. (2011). Preservice teacher education students' epistemological beliefs and conceptions about learning. *Instructional Science*, 39 (1), 87-108.
- Choy, D., Wong, A. F., Goh, K. C. & Ling Low, E. (2014). Practicum experience. Pre-service teachers' self-perception of their professional growth. *Innovations in Education and Teaching International*, 51 (5), 472-482.
- Christmann, U. (2006). Textverstehen. In J. Funke & P. Frensch (Hrsg.), *Handwörterbuch Allgemeine Psychologie: Kognition* (S. 612-620). Göttingen: Hogrefe.

- Christmann, U. & Groeben, N. (1999). Psychologie des Lesens. In B. Franzmann, K. Hasemann, D. Löffler & E. Schön (Hrsg.), *Handbuch Lesen* (S. 145-223). München: Saur. Zugriff am 08.06.2016.
- Cochran, K. F., DeRuiter, J. A. & King, R. A. (1993). Pedagogical Content Knowing: An Integrative Model for Teacher Preparation. *Journal of Teacher Education*, 44 (4), 263-272.
- Cochran, K. F., King, R. A. & DeRuiter, J. A. (Hrsg.). (1991). *Pedagogical Content Knowledge: A Tentative Model for Teacher Preparation*. Zugriff am 04.04.2016.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2. ed.). Hillsdale, NJ: Erlbaum.
- Cole, D. A., Perkins, C. E. & Zelkowitz, R. L. (2015). Impact of Homogeneous and Heterogeneous Parceling Strategies When Latent Variables Represent Multidimensional Constructs. *Psychological Methods*.
- Collin, K. & Tynjälä, P. (2003). Integrating theory and practice? Employees' and students' experiences of learning at work. *Journal of Workplace Learning*, 15 (7/8), 338-344.
- Cook, K. F., Dodd, B. G. & Fitzpatrick, S. J. (1999). A Comparison of Three Polytomous Item Response Theory Models in the Context of Testlet Scoring. *Journal of Outcome Measurement*, 3 (1), 1-20.
- Deci, E. L. & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior* (Perspectives in social psychology). New York: Plenum.
- Dehrmann, M.-G. & Standke, J. (2012). Einleitung. *Mitteilungen des Deutschen Germanistenverbandes* (2), 107-117.
- DeJong, T. & Ferguson, M. G. M. (1996). Types and qualities of knowledge. *Educational Psychologist*, 31 (2), 105-113.
- Depaepe, F., Verschaffel, L. & Kelchtermans, G. (2013). Pedagogical content knowledge: A systematic review of the way in which the concept has pervaded mathematics educational research. *Teaching and Teacher Education*, 34, 12-25.

- Desimone, L. M. (2009). Improving Impact Studies of Teachers' Professional Development. Toward Better Conceptualizations and Measures. *Educational Researcher*, 38 (3), 181-199.
- Dreyfus, S. E. (2004). The Five-Stage Model of Adult Skill Acquisition. *Bulletin of Science, Technology and Society*, 24 (3), 177-181.
- Dubberke, T., Kunter, M., McElvany, N., Brunner, M. & Baumert, J. (2008). Lerntheoretische Überzeugungen von Mathematiklehrkräften. *Zeitschrift für Pädagogische Psychologie*, 22 (34), 193-206.
- Eckes, T. (2015). Lokale Abhängigkeit von Items im TestDaF-Leseverstehen. *Diagnostica*, 61 (2), 93-106.
- Evens, M., Elen, J. & Depaepe, F. (2015). Developing Pedagogical Content Knowledge. Lessons Learned from Intervention Studies. *Education Research International*, 2015 (8), 1-23.
- Fend, H. (2008). *Schule gestalten. Systemsteuerung, Schulentwicklung und Unterrichtsqualität* (Lehrbuch, 1. Aufl.). Wiesbaden: Verl. für Sozialwiss.
- Fenstermacher, G. D. (1994). The Knower and the Known: The Nature of Knowledge in Research on Teaching. *Review of Research in Education*, 20, 3-56.
- Fernández-Balboa, J.-M. & Stiehl, J. (1995). The Generic Nature of Pedagogical Content Knowledge among College Professors. *Teaching and Teacher Education*, 11 (3), 293-306.
- Frenzel, A. C., Goetz, T., Lüdtke, O., Pekrun, R. & Sutton, R. E. (2009). Emotional transmission in the classroom. Exploring the relationship between teacher and student enjoyment. *Journal of Educational Psychology*, 101 (3), 705-716.
- Frey, A. (2004). Die Kompetenzstruktur von Studierenden des Lehrerberufs. Eine internationale Studie. *Zeitschrift für Pädagogik*, 50, 903-925.
- Friedrichsen, P. J., Abell, S. K., Pareja, E. M., Brown, P. L., Lankford, D. M. & Volkmann, M. J. (2008). Does teaching experience matter? Examining biology teachers' prior knowledge for teaching in an alternative certification program. *Journal of Research in Science Teaching*, 46 (4), 357-383.

- Fuchs, L. S., Fuchs, D., Hosp, M. K. & Jenkins, J. R. (2001). Oral Reading Fluency as an Indicator of Reading Competence: A Theoretical, Empirical, and Historical Analysis. *Scientific Studies of Reading*, 5, 239-256.
- Garbe, C. (2007). Lesen - Sozialisation - Geschlecht. Geschlechterdifferenzierende Leseforschung und -förderung. In A. Bertschi-Kaufmann (Hrsg.), *Lesekompetenz - Leseleistung - Leseförderung. Grundlagen, Modelle und Materialien* (Lehren lernen, 1. Aufl, S. 66-82). Seelze-Velber: Klett, Kallmeyer.
- Garbe, C., Holle, K. & Salisch, M. v. (2006). Entwicklung und Curriculum: Grundlagen einer Sequenzierung von Lehr-Lernzielen im Bereich des (literarischen) Lesens. In N. Groeben & B. Hurrelmann (Hrsg.), *Empirische Unterrichtsforschung in der Literatur- und Lesedidaktik. Ein Weiterbildungsprogramm* (Lesesozialisation und Medien, S. 115-154). Weinheim: Juventa-Verl.
- Gartmeier, M., Bauer, J., Fischer, M. R., Hoppe-Seyler, T., Karsten, G., Kiessling, C. et al. (2015). Fostering professional communication skills of future physicians and teachers. Effects of e-learning with video cases and role-play. *Instructional Science*, 43 (4), 443-462.
- Gerlitz, J. Y. & Schupp, J. (2005). Zur Erhebung der Big-Five-basierten Persönlichkeitsmerkmale im SOEP. Dokumentation der Instrumentenentwicklung BFI-S auf Basis des SOEP-Pretests 2005. *DIW Research, Notes 4*.
- Gold, A., Nix, D., Rieckmann, C. & Rosebrock, C. (2010). Bedingungen des Textverstehens bei leseschwachen Zwölfjährigen mit und ohne Zuwanderungshintergrund. *Didaktik Deutsch*, 28, 59-74.
- Graesser, A. C., Singer, M. & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, 101 (3), 371-395.
- Grossman, P. L. (1990). *The making of a teacher. Teacher knowledge and teacher education* (Professional development and practice series). New York: Teachers College Press.
- Grossman, P. L. & Stodolsky, S. S. (1995). Content as Context. The Role of School Subjects in Secondary School Teaching. *Educational Researcher*, 24 (8), 5-23.

- Hallsten, L. (1993). Burning out: A framework. In W. B. Schaufeli, T. Marek & C. Maslach (Hrsg.), *Professional burnout: Recent developments in theory and research* (Series in applied psychology, S. 95-113). Washington DC u.a.: Taylor & Francis.
- Hanfstingel, B. & Mayr, J. (2007). Prognose der Bewährung im Lehrerstudium und im Lehrberuf. *Journal für Lehrerinnen- und Lehrerbildung* (2), 48-55.
- Hanushek, E. A. & Rivkin, S. G. (2010). Generalizations about Using Value-Added Measures of Teacher Quality. *American Economic Review*, 100 (2), 267-271.
- Harackiewicz, J. M., Barron, K. E., Tauer, J. M. & Elliot, A. J. (2002). Predicting success in college. A longitudinal study of achievement goals and ability measures as predictors of interest and performance from freshman year through graduation. *Journal of Educational Psychology*, 94 (3), 562-575.
- Harnischfeger, A. & Wiley, D. E. (1977). Kernkonzepte des Schullernens. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 9, 207-228.
- Hascher, T. (2006). Veränderungen im Praktikum - Veränderungen durch das Praktikum. Eine empirische Untersuchung zur Wirkung von schulpraktischen Studien in der Lehrerbildung. *Zeitschrift für Pädagogik, Beiheft: 51*, 130-148.
- Haston, W. & Leon-Guerrero, A. (2008). Sources of Pedagogical Content Knowledge. Reports by Preservice Instrumental Music Teachers. *Journal of Music Teacher Education*, 17 (2), 48-59.
- Hattie, J. A. C. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement* (Reprinted.). London: Routledge.
- Hedtke, R. (2000). Das unstillbare Verlangen nach Praxisbezug. Zum Theorie-Praxis-Problem der Lehrerbildung am Exempel Schul praktischer Studien. In H.-J. Schlösser (Hrsg.), *Beruforientierung und Arbeitsmarkt* (S. 67-91). Bergisch Gladbach: Thomas Hobein.
- Heinz, M. (2013). Tomorrow's teachers—selecting the best: An exploration of the quality rationale behind academic and experiential selection criteria for initial

- teacher education programmes. *Educational Assessment, Evaluation and Accountability*, 25 (2), 93-114.
- Heinzer, S. & Oser, F. (2013). Das Advokatorische Messverfahren: Die stellvertretende Art Kompetenzen zu messen. In F. Oser, T. Bauder, P. Salzmann & S. Heinzer (Hrsg.), *Ohne Kompetenz keine Qualität. Entwickeln und Einschätzen von Kompetenzprofilen bei Lehrpersonen und Berufsbildungsverantwortlichen* (S. 139-168). Bad Heilbrunn: Klinkhardt-Verlag.
- Helmke, A., Hosenfeld, I. & Schrader, F.-W. (2004). Vergleichsarbeiten als Instrument zur Verbesserung der Diagnosekompetenz von Lehrkräften. In R. Arnold & C. Griesse (Hrsg.), *Schulleitung und Schulentwicklung* (S. 119-144). Hohengehren: Schneider-Verlag.
- Helsper, W. (1995). Pädagogisches Handeln in den Antinomien der Moderne. In H.-H. Krüger & W. Helsper (Hrsg.), *Einführung in die Grundbegriffe und Grundfragen der Erziehungswissenschaft* (S. 15-34). Opladen/Farmington Hills: Budrich.
- Helsper, W. (2014). Lehrerprofessionalität - der strukturtheoretische Professionsansatz zum Lehrerberuf. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (2. Auflage, S. 216-240). Münster u.a: Waxmann.
- Henze, I., van Driel, J. H. & Verloop, N. (2008). Development of Experienced Science Teachers' Pedagogical Content Knowledge of Models of the Solar System and the Universe. *International Journal of Science Education*, 30 (10), 1321-1342.
- Herbst, P. & Kosko, K. W. (2014). Using representations of practice to elicit mathematics teachers' tacit knowledge of practice. A comparison of responses to animations and videos. *Journal of Mathematics Teacher Education*, 17 (6), 515-537.
- Herzmann, P. & König, J. (2016). *Lehrerberuf und Lehrerbildung*. Bad Heilbrunn: Klinkhardt.
- Herzog, S. (2014). Über den Berufseinstieg hinaus: Berufsbiographien von Lehrerinnen und Lehrern im Blickfeld der Forschung. In E. Terhart (Hrsg.), *Hand-*

- buch der Forschung zum Lehrerberuf* (2., überarb. und erw. Aufl., S. 408-432). Münster: Waxmann.
- Hill, H. C., Blunk, M. L., Charalambous, C. Y., Lewis, J. M., Phelps, G. C., Sleep, L. et al. (2008). Mathematical Knowledge for Teaching and the Mathematical Quality of Instruction. An Exploratory Study. *Cognition and Instruction*, 26 (4), 430-511.
- Hill, H. C., Rowan, B. & Ball, D. L. (2005). Effects of Teachers' Mathematical Knowledge for Teaching on Student Achievement. *American Educational Research Journal*, 42, 371-406.
- Hill, H. C., Schilling, S. G. & Ball, D. L. (2004). Developing Measures of Teachers' Mathematics Knowledge for Teaching. *The Elementary School Journal*, 105 (1), 11-30.
- Hohn, K., Schiepe-Tiska, A., Sälzer, C. & Artelt, C. (2013). Lesekompetenz in PISA 2012: Veränderungen und Perspektiven. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Hrsg.), *Pisa 2012: Fortschritte und Herausforderungen in Deutschland* (Fortschritte und Herausforderungen in Deutschland). Münster: Waxmann.
- Hollins, E. R. (2011). Teacher Preparation For Quality Teaching. *Journal of Teacher Education*, 62 (4), 395-407.
- Holzberger, D., Philipp, A. & Kunter, M. (2014). Predicting teachers' instructional behaviors. The interplay between self-efficacy and intrinsic needs. *Contemporary Educational Psychology*, 39 (2), 100-111.
- Hu, L.-t. & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, 3, 424-453.
- Hülshager, U. R. & Maier, G. W. (2008). Persönlichkeitseigenschaften, Intelligenz und Erfolg im Beruf. *Psychologische Rundschau*, 59 (2), 108-122.
- Hurrelmann, B. & Groeben, N. (2006). Einleitung: Die Grundkonzeption des Weiterbildungsprogramms. In N. Groeben & B. Hurrelmann (Hrsg.), *Empirische Unterrichtsforschung in der Literatur- und Lesedidaktik. Ein Weiterbil-*

- dungsprogramm* (Lesesozialisation und Medien, S. 11-31). Weinheim: Juventa-Verl.
- IBM Corporation. (2013) IBM SPSS Statistics for Windows, Version 22.0 [Computer software]. Armonk, NY: IBM Corp.
- Iribarra, D. T. & Freund, R. (2015). *Package 'WrightMap'. Wright Map: IRT item-person map with ConQuest integration*. Zugriff am 15.03.2016. Verfügbar unter <https://cran.r-project.org/web/packages/WrightMap/WrightMap.pdf>
- Jenkins, J. R., Fuchs, L. S., van den Broek, P., Espin, C. & Deno, S. L. (2003). Sources of Individual Differences in Reading Comprehension and Reading Fluency. *Journal of Educational Psychology*, 95 (4), 719-729.
- Johnson, W. B. & Ridley, C. R. (2008). *The elements of mentoring* (2. ed., rev. ed.). Basingstoke: Palgrave Macmillan.
- Jöreskog, K. & Sörbom, D. (1993). *Structural equation modeling with the SIMPLIS command language*. Chicago: Scientific Software.
- Jütter, M. & Neuhaus, B. J. (2013). Das Professionswissen von Biologielehrkräften– Ein Vergleich zwischen Biologielehrkräften, Biologen und Pädagogen. *Zeitschrift für Didaktik der Naturwissenschaften*, 19, 31-49.
- Keller, M. M., Goetz, T., Becker, E. S., Morger, V. & Hensley, L. (2014). Feeling and showing. A new conceptualization of dispositional teacher enthusiasm and its relation to students' interest. *Learning and Instruction*, 33, 29-38.
- Kersting, M., Althoff, K. & Jäger, A. O. (2008). *WIT-2 - Wilde-Intelligenz-Test 2 (Wilde Intelligenztest 2; Wilde-Intelligenz-Test - Revision)*. PSYINDEX Tests Review. Göttingen: Hogrefe.
- Kiefer, T., Robitzsch, A. & Wu, M. (2016). *Package 'TAM'*. Zugriff am 15.03.2016. Verfügbar unter <https://cran.r-project.org/web/packages/TAM/TAM.pdf>
- Kim, T. & Schallert, D. L. (2014). Mediating effects of teacher enthusiasm and peer enthusiasm on students' interest in the college classroom. *Contemporary Educational Psychology*, 39 (2), 134-144.

- Kind, V. (2009). Pedagogical content knowledge in science education: perspectives and potential for Progress. *Studies in Science Education* (45), 169-204.
- Kintsch, W. (1992). A cognitive architecture for comprehension. In H. L. Pick, P. van den Broek & D. C. Knill (Hrsg.), *Cognition: Conceptual and methodological issues* (S. 143-163). Washington, D.C.: American Psychological Association.
- Kintsch, W. (1996). Lernen aus Texten. In J. Hoffmann & W. Kintsch (Hrsg.), *Lernen. Enzyklopädie der Psychologie. Themenbereich C, Serie II, Band 7* (S. 503-528). Göttingen: Hogrefe.
- Kintsch, W. & van Dijk, T. (1978). Toward a model of text comprehension and production. *Psychological Review* (85).
- Kirschner, S. (2013). *Modellierung und Analyse des Professionswissens von Physiklehrkräften*. Berlin: Logos.
- Kleickmann, T., Richter, D., Kunter, M., Elsner, J., Besser, M., Krauss, S. et al. (2013). Teachers' Content Knowledge and Pedagogical Content Knowledge: The Role of Structural Differences in Teacher Education. *Journal of Teacher Education*, 64 (1), 90-106.
- Kleinknecht, M. & Schneider, J. (2013). What do teachers think and feel when analyzing videos of themselves and other teachers teaching? *Teaching and Teacher Education*, 33, 13-23.
- Klicpera, C. & Gasteiger-Klicpera, B. (1993). *Lesen und Schreiben. Entwicklung und Schwierigkeiten ; die Wiener Längsschnittuntersuchungen über die Entwicklung, den Verlauf und die Ursachen von Lese- und Schreibschwierigkeiten in der Pflichtschulzeit* (Huber, 1. Aufl.). Bern: Huber.
- Klieme, E., Artelt, C., Hartig, J., Jude, N., Köller, O., Prenzel, M. et al. (Hrsg.). (2010). *PISA 2009. Bilanz nach einem Jahrzehnt*. Münster: Waxmann.
- Kliwer, H.-J. & Pohl, I. (Hrsg.). (2006). *Lexikon Deutschdidaktik*. Baltmannsweiler: Schneider-Verl. Hohengehren.
- Klusmann, U. (2011). Allgemeine berufliche Motivation und Selbstregulation. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand

- (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 277-294). Münster [u.a.]: Waxmann.
- Köller, O., Baumert, J. & Neubrand, J. (2000). Epistemologische Überzeugungen und Fachverständnis im Mathematik- und Physikunterricht. In J. Baumert, W. Bos & R. Lehmann (Hrsg.), *Mathematische und physikalische Kompetenzen am Ende der gymnasialen Oberstufe* (TIMSS/III, dritte internationale Mathematik- und Naturwissenschaftsstudie; mathematische und naturwissenschaftliche Bildung am Ende der Schullaufbahn / Jürgen Baumert ... (Hrsg.) ; Bd. 2, S. 229-269). Opladen: Leske + Budrich.
- König, J. (2007). *Welche Merkmale sollte eine „gute“ Lehrkraft haben? Gruppendiskussionen mit Schülerinnen und Schülern der zehnten Klasse*, Humboldt-Universität. Verfügbar unter <http://edoc.hu-berlin.de>
- König, J. (2015). Kontextualisierte Erfassung von Lehrerkompetenzen. *Zeitschrift für Pädagogik*, 61 (3), 305-309.
- König, J. & Blömeke, S. (2010). *Pädagogisches Unterrichtswissen. Dokumentation der Kurzfassung des TEDS-M Testinstruments zur Kompetenzmessung in der ersten Phase der Lehrerbildung*: Humboldt-Universität zu Berlin.
- König, J., Blömeke, S. & Schwippert, K. (2013). Pädagogisches Professionswissen im Studienverlauf. In S. Blömeke, A. Bremerich-Vos, G. Kaiser, G. Nold, H. Haudeck, J.-U. Keßler et al. (Hrsg.), *Professionelle Kompetenzen im Studienverlauf Weitere Ergebnisse zur Deutsch-, Englisch- und Mathematiklehrerbildung aus TEDS-LT* (S. 145-166). Münster [u.a.]: Waxmann.
- König, J., Buchholtz, C. & Dohmen, D. (2015). Analyse von schriftlichen Unterrichtsplanungen. Empirische Befunde zur didaktischen Adaptivität als Aspekt der Planungskompetenz angehender Lehrkräfte. *Zeitschrift für Erziehungswissenschaft*, 18 (2), 375-404.
- König, J., Rothland, M., Darge, K., Lünemann, M. & Tachtsoglou, S. (2013). Erfassung und Struktur berufswahlrelevanter Faktoren für die Lehrerbildung und den Lehrerberuf in Deutschland, Österreich und der Schweiz. *Zeitschrift für Erziehungswissenschaft*, 16 (3), 553-577.

- König, J. & Seifert, A. (Hrsg.). (2012). *Lehramtsstudierende erwerben pädagogisches Professionswissen. Ergebnisse der Längsschnittstudie LEK zur Wirksamkeit der erziehungswissenschaftlichen Lehrerausbildung*. Münster [u.a.]: Waxmann.
- König, J., Tachtsoglou, S., Darge, K. & Lünemann, M. (2014). Zur Nutzung von Praxis. Modellierung und Validierung lernprozessbezogener Tätigkeiten von angehenden Lehrkräften im Rahmen ihrer schulpraktischen Ausbildung. *Zeitschrift für Bildungsforschung*, 4 (1), 3-22.
- König, J., Tachtsoglou, S. & Seifert, A. (2012). Individuelle Voraussetzungen, Lerngelegenheiten und der Erwerb von pädagogischem Professionswissen. In J. König & A. Seifert (Hrsg.), *Lehramtsstudierende erwerben pädagogisches Professionswissen. Ergebnisse der Längsschnittstudie LEK zur Wirksamkeit der erziehungswissenschaftlichen Lehrerausbildung* (S. 234-283). Münster [u.a.]: Waxmann.
- Korkmaz, S., Goksuluk, D. & Zararsiz, G. (2015). *Package 'MVN'*. Zugriff am 30.03.2016. Verfügbar unter <https://cran.r-project.org/web/packages/MVN/MVN.pdf>
- Krauss, S., Blum, W., Neubrand, M., Baumert, J., Kunter, M., Besser, M. et al. (2011). Konzeptualisierung und Testkonstruktion zum fachbezogenen Professionswissen von Mathematiklehrkräften. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 135-161). Münster [u.a.]: Waxmann.
- Krauss, S. & Bruckmeier, G. (2014). Das Experten-Paradigma in der Forschung zum Lehrerberuf. In E. Terhart (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (2., überarb. und erw. Aufl., S. 241-261). Münster: Waxmann.
- Krauss, S., Neubrand, M., Blum, W., Baumert, J., Brunner, M., Kunter, M. et al. (2008). Die Untersuchung des professionellen Wissens deutscher Mathematik-Lehrerinnen und -Lehrer im Rahmen der COACTIV-Studie. *Journal für Mathematik-Didaktik*, 29 (3-4), 233-258.

- Kultusministerkonferenz. (2014). *Standards Lehrerbildung KMK-Arbeitsgruppe*. Verfügbar unter http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Standards-Lehrerbildung.pdf
- Kunina-Habenicht, O., Schulze-Stocker, F., Kunter, M., Baumert, J., Leutner, D., Förster, D. et al. (2013). Die Bedeutung der Lerngelegenheiten im Lehramtsstudium und deren individuelle Nutzung für den Aufbau des bildungswissenschaftlichen Wissens. *Zeitschrift für Pädagogik*, 59 (1), 1-24.
- Kunter, M. (2011). Motivation als Teil der professionellen Kompetenz - Forschungsbefunde zum Enthusiasmus von Lehrkräften. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 259-273). Münster [u.a.]: Waxmann.
- Kunter, M. (2014). Forschung zur Lehrermotivation. In E. Terhart, H. Bennewitz & M. Rothland (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (2. Auflage, S. 527-539). Münster u.a: Waxmann.
- Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S. & Neubrand, M. (Hrsg.). (2011). *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV*. Münster [u.a.]: Waxmann.
- Kunter, M., Kleickmann, T., Klusmann, U. & Richter, D. (2011). Die Entwicklung professioneller Kompetenz von Lehrkräften. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (S. 55-68). Münster [u.a.]: Waxmann.
- Kunter, M. & Klusmann, U. (2010). Kompetenzmessung bei Lehrkräften - Methodische Herausforderungen. *Unterrichtswissenschaft*, 38, 68-86. Measuring Teacher's Competence - Methodological Challenges.
- Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T. & Hachfeld, A. (2013). Professional competence of teachers. Effects on instructional quality and student development. *Journal of Educational Psychology*, 105 (3), 805-820.

- Kunter, M. & Trautwein, U. (2013). *Psychologie des Unterrichts* (StandardWissen Lehramt, 1., neue Ausg). Paderborn: UTB; Schöningh Paderborn.
- Kunze, I. (2004). *Konzepte von Deutschunterricht. Eine Studie zu individuellen didaktischen Theorien von Lehrerinnen und Lehrern* (Studien zur Bildungsgangforschung, Bd. 1, 1. Aufl.). Univ., FB Erziehungswiss., Habil.-Schr.--Hamburg, 2002. Wiesbaden: VS Verl. für Sozialwiss.
- LaBerge, D. & Samuels, S. J. (1974). Toward a Theory of Automatic Information Processing in Reading. *Cognitive Psychology*, 6, 293-323.
- Lee, E., Brown, M. N., Luft, J. A. & Roehrig, G. H. (2007). Assessing Beginning Secondary Science Teachers' PCK. Pilot Year Results. *School Science and Mathematics*, 107 (2), 52-60.
- Lee, E. & Luft, J. A. (2008). Experienced Secondary Science Teachers' Representation of Pedagogical Content Knowledge. *International Journal of Science Education*, 30, 1343-1363.
- Lenhard, W., Baier, H., Endlich, D., Schneider, W. & Hoffmann, J. (2013). Rethinking strategy instruction. Direct reading strategy instruction versus computer-based guided practice. *Journal of Research in Reading*, 36 (2), 223-240.
- Lepper, M. R., Corpus, J. H. & Iyengar, S. S. (2005). Intrinsic and Extrinsic Motivational Orientations in the Classroom. Age Differences and Academic Correlates. *Journal of Educational Psychology*, 97 (2), 184-196.
- Lienert, G. A. & Raatz, U. (1998). *Testaufbau und Testanalyse* (6. Auflage). Weinheim: Beltz.
- Lin, T.-C., Tsai, C.-C., Chai, C. S. & Lee, M.-H. (2013). Identifying Science Teachers' Perceptions of Technological Pedagogical and Content Knowledge (TPACK). *Journal of Science Education and Technology*, 22 (3), 325-336.
- Lindauer, T. & Schneider, H. (2007). Lesekompetenz ermitteln: Aufgaben im Unterricht. In A. Bertschi-Kaufmann (Hrsg.), *Lesekompetenz - Leseleistung - Leseförderung. Grundlagen, Modelle und Materialien* (Lehren lernen, 1. Aufl, S. 109-125). Seelze-Velber: Klett, Kallmeyer.

- Lindmeier, A. M. (2012). Video-Vignettenbasierte standardisierte Erhebung von Lehrerkognitionen. In U. Riegel & K. Macha (Hrsg.), *Fachdidaktische Forschungen* (S. 45-62). Münster: Waxmann.
- Lipowsky, F. (2006). Auf den Lehrer kommt es an. Empirische Evidenzen für Zusammenhänge zwischen Lehrerkompetenzen, Lehrerhandeln und dem Lernen der Schüler, *Beiheft 51*, 47-70.
- Little, R. J. A. (1988). A Test of Missing Completely at Random for Multivariate Data with Missing Values. *Journal of the American Statistical Association*, 83 (404), 1198-1202.
- Long, J. F. & Hoy, A. W. (2006). Interested instructors. A composite portrait of individual differences and effectiveness. *Teaching and Teacher Education*, 22 (3), 303-314.
- Magnusson, S., Krajacik, J. & Borko, H. (1999). Nature, Sources, and Development of Pedagogical Content Knowledge for Science Teaching. In J. Gess-Newsome & N. G. Lederman (Hrsg.), *Examining pedagogical content knowledge. The construct and its implications for science education* (Science & technology education library, Bd. 6, S. 95-132). Dordrecht [u.a.]: Kluwer Academic Publ.
- Mair, P., Hatzinger, R., Maier, M. & Rusch, T. (2015). *Package 'eRm'. Extended Rasch Modeling*. Zugriff am 15.03.2016. Verfügbar unter <https://cran.r-project.org/web/packages/eRm/eRm.pdf>
- Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, 57 (3), 519-530.
- Mardia, K. V. (1974). Applications of some measures of multivariate skewness and kurtosis for testing normality and robustness studies. *Sankhyā A*, 36, 115-128.
- Marsh, H. W., Hau, K.-T., Balla, J. R. & Grayson, D. (1998). Is More Ever Too Much? The Number of Indicators per Factor in Confirmatory Factor Analysis. *Multivariate behavioral research*, 33 (2), 181-220.

- Marsh, H. W. & Martin, A. J. (2011). Academic self-concept and academic achievement. Relations and causal ordering. *British Journal of Educational Psychology*, 81 (1), 59-77.
- Martin, S. D. & Dismuke, S. (2015). Teacher candidates' perceptions of their learning and engagement in a writing methods course. *Teaching and Teacher Education*, 46, 104-114.
- Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47, 149-174.
- Masters, G. N. & Wright, B. D. (1997). The partial credit mod. In van der Linden, Wim J. & R. K. Hambleton (Hrsg.), *Handbook of modern item response theory* (S. 101-121). New York: Springer.
- Mayr, J. (2014). Der Persönlichkeitsansatz in der Forschung zum Lehrberuf. In E. Terhart (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (2., überarb. und erw. Aufl., S. 189-215). Münster: Waxmann.
- Mayr, J. & Neuweg, G. H. (2006). Der Persönlichkeitsansatz in der Lehrer/innen/forschung. Grundsätzliche Überlegungen, exemplarische Befunde und Implikationen für die Lehrer/innen/bildung. In U. Greiner & M. Heinrich (Hrsg.), *Schauen, was rauskommt. Kompetenzförderung, Evaluation und Systemsteuerung im Bildungswesen* (S. 183-206). Wien: Lit.
- McElvany, N. & Schneider, C. (2009). Förderung von Lesekompetenz. In W. Lenhard (Hrsg.), *Diagnostik und Förderung des Leseverständnisses*. Göttingen: Hogrefe.
- Merzyn, G. (2002). *Stimmen zur Lehrerausbildung. Überblick über die Diskussion zur Gymnasiallehrerausbildung, basierend vor allem auf Stellungnahmen von Wissenschafts- und Bildungsgremien sowie auf Erfahrungen von Referendaren und Lehrern*. Baltmannsweiler: Schneider-Verl. Hohengehren.
- Meschede, N. (2014). *Professionelle Wahrnehmung der inhaltlichen Strukturierung im naturwissenschaftlichen Grundschulunterricht. Theoretische Beschreibung und empirische Erfassung* (Studien zum Physik- und Chemielerennen, Bd. 163). Univ., Diss.--Kiel, 2013. Berlin: Logos-Verl.

- Meschede, N., Steffensky, M., Wolters, M. & Möller, K. (2015). Professionelle Wahrnehmung der Lernunterstützung im naturwissenschaftlichen Grundschulunterricht: Theoretische Beschreibung und empirische Erfassung. *Unterrichtswissenschaft*, 43 (4), 317-335.
- Midgley, C., Feldlaufer, H. & Eccles, J. (1989). Change in teacher self efficacy and student- and taskrelated beliefs in mathematics during the transition to junior high school. *Journal of Educational Psychology*, 81 (2), 247-258.
- Mirliss, D. (2014). Preparing Future Teachers. *International Journal of Gaming and Computer-Mediated Simulations*, 6 (2), 15-29.
- Möller, J. & Schiefele, U. (2004). Motivationale Grundlagen der Lesekompetenz. In U. Schiefele, C. Artelt, W. Schneider & P. Stanat (Hrsg.), *Struktur, Entwicklung und Förderung von Lesekompetenz. Vertiefende Analysen im Rahmen von PISA 2000* (1. Aufl, S. 101-124). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Möller, J. & Trautwein, U. (2015). Selbstkonzept. In E. Wild & J. Möller (Hrsg.), *Pädagogische Psychologie* (Springer-Lehrbuch, 2. Aufl. 2015. vollst. überarb. u. aktualisierte, S. 177-200). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Moosbrugger, H. & Kelava, A. (2012). *Testtheorie und Fragebogenkonstruktion* (Springer-Lehrbuch, 2., aktualisierte und überarbeitete Auflage). Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg.
- Mucherah, W. & Ambrose-Stahl, D. (2014). Relation of reading motivation to reading achievement in seventh-grade students from Kenya and the United States. *International Perspectives in Psychology: Research, Practice, Consultation*, 3 (3), 154-166.
- Muraki, E. (1993). Information Functions of the Generalized Partial Credit Model. *Applied Psychological Measurement*, 17 (4), 351-363.
- Muthén, L. K. & Muthén, B. O. (2012). *Mplus User's Guide* (Seventh Edition). Los Angeles, CA: Muthén & Muthén.
- Myers, N. D., Ahn, S. & Jin, Y. (2011). Sample Size and Power Estimates for a Confirmatory Factor Analytic Model in Exercise and Sport: A Monte Carlo Approach. *Research Quarterly for Exercise and Sport*, 82 (3), 412-423.

- Narciss, S. & Proske, A. (2001). *Analyse und Beschreibung von Lern- und Studienaufgaben für die universitäre Lehre. Unveröffentlichtes Manuskript*, TU Dresden.
- Naumann, J., Artelt, C., Schneider, W. & Stanat, P. (2010). Lesekompetenz von PISA 2000 bis PISA 2009. In E. Klieme, C. Artelt, J. Hartig, N. Jude, O. Köller, M. Prenzel et al. (Hrsg.), *PISA 2009. Bilanz nach einem Jahrzehnt* (S. 23-66). Münster: Waxmann.
- Neisser, U., Boodoo, G., Bouchard Jr., Thomas J., Boykin, A. W., Brody, N. et al. (1996). Intelligence: Knowns and Unknowns. *American Psychologist*, 51 (2), 77-101.
- Neugebauer, M. (2013). Wer entscheidet sich für ein Lehramtsstudium – und warum? Eine empirische Überprüfung der These von der Negativselektion in den Lehrerberuf. *Zeitschrift für Erziehungswissenschaft*, 16 (1), 157-184.
- Neuweg, G. H. (2005). Wie grau ist alle Theorie, wie grün des Lebens goldner Baum? LehrerInnenausbildung im Spannungsfeld von Theorie und Praxis. *ÖFEB-Newsletter*, 5 (1), 5-15.
- Neuweg, G. H. (2014). Das Wissen der Wissensvermittler. In E. Terhart (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (2., überarb. und erw. Aufl., S. 583-614). Münster: Waxmann.
- Neuweg, G. H. (2015). Kontextualisierte Kompetenzmessung. Eine Bilanz zu aktuellen Konzeptionen und forschungsmethodischen Zugängen. *Zeitschrift für Pädagogik*, 61, 377-383.
- Nix, D. (2011). *Förderung der Leseflüssigkeit. Theoretische Fundierung und empirische Überprüfung eines kooperativen Lautlese-Verfahrens im Deutschunterricht* (Lesesozialisation und Medien). Univ., Diss.--Frankfurt am Main, 2010. Weinheim: Juventa-Verl.
- North, K., Reinhardt, K. & Sieber-Suter, B. (2012). *Kompetenzmanagement in der Praxis- Mitarbeiterkompetenzen systematisch identifizieren, nutzen und entwickeln*. Wiesbaden: Gabler Verlag.
- Olgetree, G. L. (2008). The effect of fifth grade science teachers' pedagogical content knowledge on their decision making and student learning outcomes on the

- concept of chemical change. *Dissertation Abstracts International Section: Humanities and Social Sciences*, 69, 1724.
- Olszewski, J. (2010). *The impact of physics teachers' pedagogical content knowledge on teacher action and student outcomes* (Studien zum Physik- und Chemielernen, Bd. 109). Essen, Univ., Diss.--Duisburg, 2010. Berlin: Logos Verl.
- Opitz, M. F. & Rasinski, T. (1998). *Good-bye round robin. Twenty-five effective oral reading strategies*. Portsmouth, NH: Heinemann.
- Oser, F. (2006). Zugänge ermöglichen, Zugänge verwehren: Entwurf eines Ausleseverfahrens in der Lehrerbildung (ein Essay). *Beiträge zur Lehrerbildung*, 24 (1), 30-42.
- Oser, F., Curcio, G.-P. & Düggeli, A. (2007). Kompetenzmessung in der Lehrerbildung als Notwendigkeit Fragen und Zugänge. *Beiträge zur Lehrerbildung*, 25, 14-25.
- Oser, F. & Forster-Heinzer, S. (2015). Wer setzt das Maß ? Eine kritische Auseinandersetzung mit dem Advokatorischen Ansatz. *Zeitschrift für Pädagogik*, 61 (3), 361-377.
- Oser, F., Heinzer, S. & Salzmann, P. (2010). Die Messung der Qualität von professionellen Kompetenzprofilen von Lehrpersonen mit Hilfe der Einschätzung von Filmvignetten. *Unterrichtswissenschaft*, 38 (1), 5-28.
- Özden, M. (2008). The Effect of Content Knowledge on Pedagogical Content Knowledge: The Case of Teaching Phases of Matters. *Educational Sciences: Theory & Practice*, 8 (2).
- Pädagogische Hochschule Heidelberg (Hrsg.). (2016). *Zentrales Handlungsfeld Studium und Lehre*. Zugriff am 29.04.2016. Verfügbar unter http://www.ph-heidelberg.de/fileadmin/de/hochschule/Struktur-_und_Entwicklungsplan/PHHD_StEP_I._Zentrales_Handlungsfeld_Studium_und_Lehre.pdf
- Palmer, D. J., Stough, L. M., Burdenski Jr., T. K. & Gonzales, M. (2005). Identifying Teacher Expertise: An Examination of Researchers' Decision Making. *Educational Psychologist*, 40 (1), 13-25.

- Park, S. & Chen, Y.-C. (2012). Mapping out the integration of the components of pedagogical content knowledge (PCK). Examples from high school biology classrooms. *Journal of Research in Science Teaching*, 49 (7), 922-941.
- Patry, J.-L. (2000). Schulunterricht ist komplex – Kann da Theorie noch praktisch sein? *Salzburger Beiträge zur Erziehungswissenschaft*, 4 (1), 43-59.
- Paulick, I., Großschedl, J., Harms, U. & Möller, J. (2016). Preservice Teachers Professional Knowledge and Its Relation to Academic Self-Concept. *Journal of Teacher Education*.
- Pause, G. (1970). Merkmale der Lehrerpersönlichkeit. In K. Ingenkamp (Hrsg.), *Handbuch der Unterrichtsforschung* (Bd. 2, S. 1353-1529). Weinheim: Beltz.
- Pfost, M., Dörfler, T. & Artelt, C. (2010). Der Zusammenhang zwischen außerschulischem Lesen und Lesekompetenz. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 42 (3), 167-176.
- Philipp, M. (2013). Die feinen Unterschiede: Ein Forschungsüberblick zum Zusammenhang von Geschlecht und Lesemotivation sowie -verstehen. In B. Lundt & T. Tholen (Hrsg.), *"Geschlecht" in der Lehramtsausbildung. Die Beispiele Geschichte und Deutsch* (Historische Geschlechterforschung und Didaktik, Bd. 3, S. 297-321). Berlin u.a.: Lit.
- Philipp, M. & Schilcher, A. (Hrsg.). (2012). *Selbstreguliertes Lesen. Ein Überblick über wirksame Leseförderansätze* (1. Aufl.). Seelze-Velber: Klett/Kallmeyer.
- Pissarek, M. & Schilcher, A. (2015). Fachspezifische Lehrerkompetenzen im Fach Deutsch messen? Modellierung und Konstruktvalidierung eines Erhebungsinstruments im Rahmen der der Projektgruppe FALKO Regensburg. In C. R. Bräuer (Hrsg.), *Lehrende im Blick. Empirische Lehrerforschung in der Deutschdidaktik* (SpringerLink : Bücher, S. 321-342). Wiesbaden: Springer VS.
- Pohlmann, B. & Möller, J. (2010). Fragebogen zur Erfassung der Motivation für die Wahl des Lehramtsstudiums (FEMOLA). *Zeitschrift für Pädagogische Psychologie*, 24 (1), 73-84.

- Polanyi, M. (1985). *Implizites Wissen* (Bd. 543, 1. Aufl.). Frankfurt am Main: Suhrkamp.
- Rao, D. & Stupans, I. (2012). Exploring the potential of role play in higher education. Development of a typology and teacher guidelines. *Innovations in Education and Teaching International*, 49 (4), 427-436.
- Rasinski, T. V. (2005). *The fluent reader. Oral reading strategies for building word recognition, fluency, and comprehension* (Theory and practice). New York: Scholastic Professional Books.
- Rausch, T., Matthäi, J. & Artelt, C. (2015). Mit Wissen zu akkurateren Urteilen? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 47 (3), 147-158.
- Reinders & Heinz. (2006). Kausalanalysen in der Längsschnittforschung. Das Crossed-Lagged-Panel Design. *Diskurs Kindheits- und Jugendforschung*, 1 (4), 569-587.
- Reinisch, H. (2009). Lehrprofessionalität als theoretischer Term. In O. Zlatkin-Troitschanskaia (Hrsg.), *Lehrprofessionalität. Bedingungen, Genese, Wirkungen und ihre Messung* (Beltz-Bibliothek, S. 33-43). Weinheim: Beltz.
- Resnick, L. & Hall, M. (1998). Learning Organizations for Sustainable Education Reform. *Education Yesterday, Education Tomorrow* (4), 89-118.
- Retelsdorf, J., Bauer, J., Gebauer, S. K., Kauper, T. & Möller, J. (2014a). Erfassung berufsbezogener Selbstkonzepte von angehenden Lehrkräften (ERBSE-L). *Diagnostica*, 60 (2), 98-110.
- Retelsdorf, J., Bauer, J., Gebauer, S. K., Kauper, T. & Möller, J. (2014b). Erfassung berufsbezogener Selbstkonzepte von angehenden Lehrkräften (ERBSE-L). *Diagnostica*, 60 (2), 98-110.
- Richter, T. & Christmann, U. (2002). Lesekompetenz: Prozessebenen und interindividuelle Unterschiede. In N. Groeben & B. Hurrelmann (Hrsg.), *Lesekompetenz. Bedingungen, Dimensionen, Funktionen* (Lesesozialisation und Medien, S. 25-58). Weinheim: Juventa-Verl.

- Riese, J. & Reinhold, P. (2008). Entwicklung und Validierung eines Instruments zur Messung professioneller Handlungskompetenz bei (angehenden) Physiklehrkräften. *Lehrerbildung auf dem Prüfstand*, 1, 625-640.
- Riss, M. (2005). Vielfalt ist gefragt! *Die Grundschulzeitschrift*, 52-55.
- Roloff Henoch, J., Klusmann, U., Lüdtke, O. & Trautwein, U. (2015). Die Entwicklung beruflicher Selbstregulation. Ein Vergleich zwischen angehenden Lehrkräften und anderen Studierenden. *Zeitschrift für Pädagogische Psychologie*, 29 (3-4), 151-162.
- Rosebrock, C. & Nix, D. (2014). *Grundlagen der Lesedidaktik* (und der systematischen schulischen Leseförderung). Baltmannsweiler: Schneider Hohengehren.
- Rosebrock, C., Nix, D., Rieckmann, C. & Gold, A. (2011). *Leseflüssigkeit fördern. Lautleseverfahren für die Primar- und Sekundarstufe*. Seelze: Klett, Kallmeyer.
- Ross, J. A. (1992). Teacher efficacy and the effect of coaching on student achievement: Results of a randomized field trial. *Canadian Journal of Education*, 17 (1), 51-65.
- Rost, J. (2004). *Lehrbuch Testtheorie, Testkonstruktion*. Bern: Hans Huber.
- Rothland, M. & Tirre, S. (2011). Selbsterkundung für angehende Lehrkräfte. Was erfassen ausgewählte Verfahren der Eignungsabklärung? *Zeitschrift für Pädagogik*, 57 (5), 655-673.
- Runge, T. E., Frey, D., Gollwitzer, P. M., Helmreich, R. L. & Spence, J. T. (1981). Masculine (instrumental) and feminine (expressive) traits: A comparison between Students in the United States and West Germany. *Journal of Cross-Cultural Psychology*, 12 (2), 142-162.
- Rupp, G. & Bonholt, H. (2006). Lehr-Lern-Forschung als empirische Lese-/Literaturdidaktik?! In N. Groeben & B. Hurrelmann (Hrsg.), *Empirische Unterrichtsforschung in der Literatur- und Lesedidaktik. Ein Weiterbildungsprogramm* (Lesesozialisation und Medien, S. 239-253). Weinheim: Juventa-Verl.
- Savery, J. R. & Duffy, T. M. (1995). Problem based learning: An instructional model and its constructivist framework. *Educational Technology*, 35, 31-38.

- Schaarschmidt, U. (Hrsg.). (2005). *Halbtagsjobber? Psychische Gesundheit im Lehrerberuf - Analyse eines veränderungsbedürftigen Zustandes* (2. Aufl., Dr. nach Typoskript). Weinheim: Beltz.
- Schaarschmidt, U. & Fischer, A. (1997). ADEM - ein diagnostisches Instrument zur Differenzierung von Typen gesundheitsrelevanten Verhaltens und Erlebens gegenüber der Arbeit. *Zeitschrift für Differenzielle und Diagnostische Psychologie*, 18 (3), 151-163.
- Scherf, D. (2013). *Leseförderung aus Lehrersicht. Eine qualitativ-empirische Untersuchung professionellen Wissens* (SpringerLink : Bücher). Wiesbaden: Springer VS.
- Schiefele, U., Artelt, C., Schneider, W. & Stanat, P. (Hrsg.). (2004). *Struktur, Entwicklung und Förderung von Lesekompetenz. Vertiefende Analysen im Rahmen von PISA 2000* (1. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Schiefele, U. & Schaffner, E. (2015). Teacher interests, mastery goals, and self-efficacy as predictors of instructional practices and student motivation. *Contemporary Educational Psychology*, 42, 159-171.
- Schleicher, A. (2012). Building a High-Quality Teaching Profession. Lessons from around the world. *Educational Studies* (1), 74-92.
- Schmelzing, S., Wüsten, S., Sandmann, A. & Neuhaus, B. J. (2008). Evaluation von zentralen Inhalten der Lehrerbildung: Ansätze zur Diagnostik des fachdidaktischen Wissens von Biologielehrkräften. *Lehrerbildung auf dem Prüfstand*, 1, 641-663.
- Schmelzing, S., Wüsten, S., Sandmann, A. & Neuhaus, B. J. (2010). Fachdidaktisches Wissen und Reflektieren im Querschnitt der Biologielehrerbildung. *Zeitschrift für Didaktik der Naturwissenschaften*, 16, 189-207.
- Schneider, W., Schlagmüller, M. & Ennemoser, M. (2007). *LGVT 6-12: Leseschwindigkeits- und -verständnistest für die Klassen 6-12*. Göttingen: Hogrefe.
- Schnotz, W. & Dutke, S. (2004). Kognitionspsychologische Grundlagen der Lesekompetenz: Mehrebenenverarbeitung anhand multipler Informationsquellen. In U. Schiefele, C. Artelt, W. Schneider & P. Stanat (Hrsg.), *Struktur, Entwick-*

- lung und Förderung von Lesekompetenz. *Vertiefende Analysen im Rahmen von PISA 2000* (1. Aufl, S. 61-99). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Schön, E. (2002). Einige Anmerkungen zur PISA -Studie, auch aus literaturdidaktischer Perspektive. Oder: Lesen lernt man nur durch Lesen. In K. Franz & F.-J. Payrhuber (Hrsg.), *Lesen Heute. Leseverhalten von Kindern und Jugendlichen und Leseförderung im Kontext der PISA -Studie* (S. 72-91). Kronach: Schneider.
- Schönbrodt, F. D. & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality*, 47 (5), 609-612.
- Schwarzer, R. & Warner, L. M. (2014). Forschung zur Selbstwirksamkeit bei Lehrerinnen und Lehrern. In E. Terhart (Hrsg.), *Handbuch der Forschung zum Lehrerberuf* (2., überarb. und erw. Aufl., S. 662-678). Münster: Waxmann.
- Seidel, T., Blomberg, G. & Stürmer, K. (2010). "Observer" - Validierung eines videobasierten Instruments zur Erfassung der professionellen Wahrnehmung von Unterricht. In E. Klieme, D. Leutner & M. Kenk (Hrsg.) *Kompetenzmodellierung. Zwischenbilanz des DFG-Schwerpunktprogramms und Perspektiven des Forschungsansatzes*, 296-306 [Themenheft]. Beltz.
- Sherin, M. G. & van Es, E. A. (2008). Effects of Video Club Participation on Teachers' Professional Vision. *Journal of Teacher Education*, 60, 20-37.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15, 4-14.
- Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57, 1-21.
- Simmons, P. E., Emory, A., Carter, T., Coker, T., Finnegan, B., Crockett, D. et al. (1999). Beginning Teachers: Beliefs and Classroom Actions. *Journal of Research in Science Teaching*, 36 (8), 930-954.
- Souvignier, E. & Mokhlesgerami, J. (2005). Implementation eines Programms zur Vermittlung von Lesestrategien im Deutschunterricht. *Zeitschrift für Pädagogische Psychologie*, 19 (4), 249-261.

- Spinner, K. H. (2004). Lesekompetenz in der Schule. In U. Schiefele, C. Artelt, W. Schneider & P. Stanat (Hrsg.), *Struktur, Entwicklung und Förderung von Lesekompetenz. Vertiefende Analysen im Rahmen von PISA 2000* (1. Aufl., S. 125-139). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Streblow, L. (2004). Zur Förderung der Lesekompetenz. In U. Schiefele, C. Artelt, W. Schneider & P. Stanat (Hrsg.), *Struktur, Entwicklung und Förderung von Lesekompetenz. Vertiefende Analysen im Rahmen von PISA 2000* (1. Aufl., S. 275-306). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Südkamp, A., Möller, J. & Pohlmann, B. (2008). Der Simulierte Klassenraum. *Zeitschrift für Pädagogische Psychologie*, 22 (3), 261-276.
- Sutter, J. (2016). *Grafische Visualisierungen bei der Stellenübergabe. Ein Werkzeug zur Externalisierung von implizitem Wissen* (Schriftenreihe der Kalaidos Fachhochschule Schweiz). Wiesbaden: Springer Gabler.
- Tepner, O., Borowski, A., Dollny, S., Fischer, H. E., Jüttner, M., Kirschner, S. et al. (2012). Modell zur Entwicklung von Testitems zur Erfassung des Professionswissens von Lehrkräften in den Naturwissenschaften. *Zeitschrift für Didaktik der Naturwissenschaften*, 18, 7-28.
- Tepner, O. & Dollny, S. (2014). Entwicklung eines Testverfahrens zur Analyse fachdidaktischen Wissens. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 311-323). Berlin, Heidelberg: Springer.
- Terhart, E. (1994). *Berufsbiographien von Lehrern und Lehrerinnen*. Frankfurt am Main: Lang.
- Terhart, E. (2007). Erfassung und Beurteilung der beruflichen Kompetenz von Lehrkräften. In M. Lüders (Hrsg.), *Forschung zur Lehrerbildung. Kompetenzentwicklung und Programmevaluation* (S. 37-62). Münster, New York: Waxmann.
- Thielsch, M. T. & Brandenburg, T. (Hrsg.). (2012). *Praxis der Wirtschaftspsychologie II. Themen und Fallbeispiele für Studium und Anwendung*. Münster: MV Wissenschaft.

- Tschannen-Moran, M., Woolfolk Hoy, A. & Hoy, W. K. (1998). Teacher Efficacy: Its meaning and measure. *Review of Educational Research*, 68 (2), 202-248.
- Uyulgan, M. & Akkuzu, N. (2014). An Overview of Student Teachers' Academic Intrinsic Motivation. *Educational Sciences: Theory & Practice*, 14 (1), 24-32.
- Van Dijk, E. M. & Kattmann, U. (2010). Evolution im Unterricht: Eine Studie über fachdidaktisches Wissen von Lehrerinnen und Lehrern. *Zeitschrift für Didaktik der Naturwissenschaften*, 16, 7-21.
- Van Driel, J., Beijaard, D. & Verloop, N. (2001). Professional development and reform in science education: The role of teachers' practical knowledge. *Journal of Research in Science Teaching*, 38 (2), 137-158.
- Van Driel, J. H. & Berry, A. (2012). Teacher Professional Development Focusing on Pedagogical Content Knowledge. *Educational Researcher*, 41 (1), 26-28.
- Van Driel, J. H., Verloop, N. & Vos, W. de. (1998). Developing Science Teachers' Pedagogical Content Knowledge. *Journal of Research in Science Teaching*, 35, 673-695.
- Vandenberghe, R. (Ed.). (2006). *Understanding and preventing teacher burnout. A sourcebook of international research and practice* (Digitally print. 1. paperback version). Cambridge: Cambridge Univ. Press.
- Vansteenkiste, M., Leny, W. & Deci, E. L. (2006). Intrinsic Versus Extrinsic Goal Contents in Self-Determination Theory: Another Look at the Quality of Academic Motivation. *Educational Psychologist*, 41 (1), 19-31.
- Veenendaal, N. J., Groen, M. A. & Verhoeven, L. (2015). What oral text reading fluency can reveal about reading comprehension. *Journal of Research in Reading*, 38 (3), 213-225.
- Vogelsang, C. & Reinhold, P. (2013). Zur Handlungsvalidität von Tests zum professionellen Wissen von Lehrkräften. *Zeitschrift für Didaktik der Naturwissenschaften*, 19, 103-128.
- Voss, T., Kleickmann, T., Kunter, M. & Hachfeld, A. (2011). Überzeugungen von Mathematiklehrkräften. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Hrsg.), *Professionelle Kompetenz von Lehrkräften*.

- Ergebnisse des Forschungsprogramms COACTIV* (S. 235-257). Münster [u.a.]: Waxmann.
- Voss, T., Kunina-Habenicht, O., Hoehne, V. & Kunter, M. (2015). Stichwort Pädagogisches Wissen von Lehrkräften. Empirische Zugänge und Befunde. *Zeitschrift für Erziehungswissenschaft*, 18 (2), 187-223.
- Wainer, H. & Kiely, G. (1987). Item Clusters and Computized Adaptive Testing: A Case for Testlets. *Journal of Educational Measurement*, 37 (3).
- Watt, H. M. G. & Richardson, P. W. (2007). Motivational Factors Influencing Teaching as a Career Choice: Development and Validation of the FIT-Choice Scale. *The Journal of Experimental Education*, 75 (3), 167-202.
- Weinert, E. F. & Helmke, A. (1996). Der gute Lehrer: Person, Funktion oder Fiktion? In A. Leschinsky (Hrsg.), *Die Institutionalisierung von Lehren und Lernen. Beiträge zu einer Theorie der Schule* (Zeitschrift für Pädagogik. Beiheft, S. 223-233). Weinheim: Beltz.
- Weinert, F. E. (1996). 'Der gute Lehrer', 'die gute Lehrerin' im Spiegel der Wissenschaft. *Beiträge zur Hochschulbildung*, 14 (2), 141-151.
- Weinert, F. E. (Hrsg.). (2001). *Leistungsmessungen in Schulen* (3. Aufl.). Weinheim und Basel: Beltz.
- Whitehead, A. N. (1929). *The aims of education and other essays*. London: Macmillan.
- Widaman, K. F. & Reise, S. P. (1997). Exploring the measurement invariance of psychological instruments: Applications in the substance use domain. In K. J. Bryant, M. Windle & S. G. West (Hrsg.), *The science of prevention: Methodological advances from alcohol and substance abuse research* (S. 281-324). Washington: American Psychological Association.
- Wieser, D. (2008). *Literaturunterricht aus Sicht der Lehrenden. Eine qualitative Interviewstudie*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Willenberg, H. (2005). Ein handhabares System, um Textschwierigkeiten einzuschätzen. Vorschläge für eine Textdatenbank von Sachtexten. In M. Fix & R. Jost (Hrsg.), *Sachtexte im Deutschunterricht*. (S. 94-107). Baltmannsweiler: Schneider Hohengehren.

- Wirtz, M. & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität. Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen*. Göttingen: Hogrefe Verl. für Psychologie.
- Witner, S. & Tepner, O. (2011). Entwicklung geschlossener Testaufgaben zur Erhebung des fachdidaktischen Wissens von Chemielehrkräften. *chimica et ceterae artes rerum naturae didacticae*, 37, 113-137.
- Wright, B. & Panchapakesan, N. (1969). A procedure for sample-free item analysis. *Educational and Psychological Measurement*, 29, 23-48.
- Wu, M. & Adams, R. (2007). Applying the Rasch Model to psycho-social Measurement: A practical Approach. *Educational Measurement Solutions*. Zugriff am 16.03.2016. Verfügbar unter [http://media.metrik.de/uploads/incoming/pub/Literatur/von%20Winfried/Rasch Measurement_Complete.pdf](http://media.metrik.de/uploads/incoming/pub/Literatur/von%20Winfried/Rasch%20Measurement_Complete.pdf)
- Yen, W. M. (1984). Effects of Local Item Dependence on the Fit and Equating Performance of the Three-Parameter Logistic Model. *Applied Psychological Measurement*, 8 (2), 125-145. Zugriff am 08.11.2016. Verfügbar unter <http://conservancy.umn.edu/bitstream/handle/11299/107543/1/v08n2p125.pdf>
- Yen, W. M. (1993). Scaling Performance Assessments: Strategies for Managing Local Item Dependence. *Journal of Educational Measurement*, 30 (3), 187-213.
- Yeung, A. S., Craven, R. G. & Kaur, G. (2014). Teachers' self-concept and valuing of learning. Relations with teaching approaches and beliefs about students. *Asia-Pacific Journal of Teacher Education*, 42 (3), 305-320.
- Zaubauer, A. C. M., Brouer, B., Schmidt, A. & Möller, J. (2015). Kleine Veränderung – großer Gewinn? Effekte struktureller Veränderungen in der gymnasialen Lehrerbildung. *Beiträge zur Hochschulbildung*, 37 (4), 36-50.
- Zwaan, R. A. & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, 123 (2), 162-185.

Anhang: Vignetten des finalen Testinstruments

Die Unterrichtsvignetten können bei der Autorin angefragt werden.